INFORMATION AND CONTROL **48**, 221–260 (1981)

Sequential, Continuous and Parallel Grammars

H. C. M. KLEIJN AND G. ROZENBERG

Institute of Applied Mathematics and Computer Science, University of Leiden, Wassenaarseweg 80, Leiden, The Netherlands

Sequential and parallel ways of rewriting are investigated and compared in the framework of selective substitution grammars. New aspects of the notion of generative determinism of a grammar and of the notion of symmetric context are studied. Several new characterizations of known classes of languages are obtained.

INTRODUCTION

Selective substitution grammars (see, Rozenberg, 1977; Rozenberg and Wood, 1980; Ehrenfeucht *et al.*, 1980; Kleijn and Rozenberg, 1980) form a general framework for rewriting systems. They allow one to study common features of seemingly different kinds of grammars as well as form a "guide" for discovering new, but natural, classes of grammars. This paper is an example of research in the first direction. One of the important, and currently active, research topics in formal language theory is the investigation of the difference between parallel and sequential rewriting. In our paper we consider this topic in the general framework of selective substitution grammars. We investigate here three classes of selective substitution grammars: sequential grammars (abbreviated S grammars), parallel grammars (abbreviated L grammars) and continuous grammars (abbreviated C grammars).

S grammars formalize the notion of a rewriting system in which *only one* occurrence of a letter is rewritten in a derivation step, L grammars formalize the notion of a rewriting system in which all occurrences of *all* letters are rewritten in a derivation step and C grammars formalize the notion of an "in-between" rewriting system, where an arbitrary number of occurrences *forming a continuous segment* is rewritten in a derivation step. (C grammars were introduced in Ehrenfeucht *et al.*, 1980.)

The paper is organized as follows.

In Section II we define formally the above-mentioned three classes of selective substitution grammars and the classes of languages they generate. (denoted $\mathscr{L}(S)$, $\mathscr{L}(L)$ and $\mathscr{L}(C)$, respectively).

In Section III we establish the position of all three classes of languages with respect to several known classes of languages (namely, the classes of

0019-9958/81/030221–40$02.00/0

context-free, $E0L$ and $ET0L$ languages). We establish several normal forms for the classes of grammars we consider and investigate the mutual relationships between $\mathscr{L}(S)$, $\mathscr{L}(L)$ and $\mathscr{L}(C)$.

In Section IV we consider two new kinds of generative determinism. Both reflect the idea of "*deterministically* fitting a word under rewriting into (a component of) the selector." It turns out that considering such a determinism allows one to point out essential differences between S grammars, L grammars and C grammars; it also yields new characterizations of context-free, linear and regular languages.

In Section V we study a (in our opinion very important) notion of "context symmetry" in rewriting systems. We provide two different formalizations of this idea in the framework of S grammars, L grammars and C grammars and investigate the effect of both "restrictions" on the language generating power of these classes of grammars.

I. Preliminaries

We assume the reader to be familiar with the basic concepts of formal language theory as, e.g., in the scope of Salomaa (1973) and Rozenberg and Salomaa (1980). Perhaps the following notations, definitions and results require an additional explanation.

For a word x, $|x|$ denotes its length, $\#_a x$ denotes the number of occurrences of a in x and *alph* x denotes the set of letters occurring in x. Λ denotes the empty word.

In the rewriting systems that we will consider that use context-free productions, those productions are defined either in the form $A \to \alpha$ or by a finite substitution h, in which case $A \to \alpha$ is denoted as $\alpha \in h(A)$, where A is a letter and α is a word.

To save unnecessary writing we often define in constructions a finite substitution by stating which productions should be in; in these cases it is implicit that we mean the smallest finite substitution satisfying all stated conditions.

If $A \to \alpha$ is a production, then it is termed an *A-production* and α is referred to as the *right-hand side* of this production.

For a rewriting system G, *maxr* (G) denotes the maximal length of the right-hand sides of all productions in G and *sent* G denotes all words that can be derived from the axiom of G. G is *propagating* if Λ does not occur as a right-hand side of any production of G, otherwise we say that G has *erasing* productions.

We consider two languages L_1 and L_2 to be equal if $L_1 \cup \{\Lambda\} = L_2 \cup \{\Lambda\}$. Two rewriting systems are *equivalent* if the languages they generate are equal.

$\mathscr{L}(\text{Reg})$, $\mathscr{L}(CF)$, $\mathscr{L}(E0L)$ and $\mathscr{L}(ET0L)$ denote the classes of languages generated by respectively, regular grammars, context-free grammars, $E0L$ systems and $ET0L$ systems. Such languages are said to be regular, context-free, $E0L$ and $ET0L$, respectively.

Let $G = (\Sigma, P, S, \Delta)$ be a context-free grammar, where Σ is the total alphabet of G, Δ is the terminal alphabet of G, $S \in \Sigma \setminus \Delta$ the axiom of G and P the set of productions of G. If $A \to \alpha \in P$, such that $A \in \Sigma \setminus \Delta$ and $\alpha \in \Delta^* \cup \Delta^*(\Sigma \setminus \Delta)$ (or $\alpha \in \Delta^* \cup (\Sigma \setminus \Delta) \Delta^*$, or $\alpha \in \Delta^* \cup \Delta^* (\Sigma \setminus \Delta)\Delta^*$) then $A \to \alpha$ is termed a *right-linear production* (respectively *left-linear production* or *linear production*).

If all $A \to \alpha \in P$ are right-linear (resp. left-linear or linear) then G is called a *right-linear* (resp. *left-linear* or *linear*) *grammar*. It is well known that the classes of languages generated by left-linear and right-linear grammars equal $\mathscr{L}(\text{Reg})$ and that the class of languages generated by linear grammars strictly contains $\mathscr{L}(\text{Reg})$ and is strictly contained in $\mathscr{L}(CF)$.

It is often convenient to extend the set of productions of a context-free grammar G such that G contains for every symbol of its total alphabet a production. The so obtained construct is referred to as an *E0S system* (*EP0S system* if it is propagating). $E0S$ systems are specified in the form (Σ, h, S, Δ), where Σ, S and Δ are as in the description of a context-free grammar and h is a finite substitution, from Σ^* into Σ^*.

It is easy to see that the class of languages generated by $E0S$ systems equals $\mathscr{L}(CF)$.

II. Basic Definitions

In this section the notions basic to this paper are defined.

Throughout the whole paper barred versions of symbols are used with a special meaning: the original symbol is *activated.* If Σ is an alphabet, then the homomorphism *iden* from $(\bar{\Sigma} \cup \Sigma)^*$ into Σ^* is defined by *iden* $\bar{a} = a$ and *iden* $a = a$.

Originally selective substitution grammars were defined by Rozenberg (1977). In this paper, however, we will use the definition of an $E0S$ based s-grammar from Kleijn and Rozenberg (1980) as a basic notion.

Definition II.1. An *E0S based s-grammar H* is a 5-tuple $(\Sigma, h, S, \Delta, K)$, where *base* $H = (\Sigma, h, S, \Delta)$ is an $E0S$ system and K, the selector of H, denoted *sel* H, is a language over $\Sigma \cup \bar{\Sigma}$.

If $v, w \in \Sigma^*$, then v *directly derives* w (*in* H), denoted $v \Rightarrow_H w$, if there exists a word $x \in K$, such that $x \neq v$, *iden* $x = v$ and if $x = a_1 ..., a_n$, with $a_i \in \Sigma \cup \bar{\Sigma}$, $1 \leqslant i \leqslant n$, then $w = \alpha_1 ... \alpha_n$, with for each i, $\alpha_i = a_i$ if $a_i \in \Sigma$ and

$a_i \in h$ (*iden* a_i) if $a_i \in \bar{\Sigma}$. Let $\Rightarrow_H^*$ denote the reflexive and transitive closure of $\Rightarrow_H$. (We use $\Rightarrow^*$ and $\Rightarrow$, respectively, if no confusion is possible.)

The *language of* H, denoted $L(H)$ is defined by $L(H) = sent\ H \cap \Delta^* = \{w \in \Delta^* : S \Rightarrow^* w\}$. ▮

The subject of investigation of this paper are $E0S$ based s-grammars which formalize sequential, continuous and parallel rewriting.

DEFINITION II.2. Let $H = (\Sigma, h, S, \Delta, K)$ be an $E0S$ based s-grammar.

(i) H is termed an *n-sequential grammar*, *nS grammar* for short, if $K = \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^*$, for some X_i, Y_i, $Z_i \subseteq \Sigma$, $1 \leqslant i \leqslant n$. H is *sequential*, (H is an *S grammar*), if it is n-sequential for some $n \geqslant 1$.

(ii) H is termed an *n-continuous grammar*, *nC grammar* for short, if $K = \bigcup_{i=1}^n X_i^* \bar{Y}_i^+ Z_i^*$, for some X_i, Y_i, $Z_i \subseteq \Sigma$, $1 \leqslant i \leqslant n$. H is *continuous* (H is a *C grammar*), if it is n-continuous for some $n \geqslant 1$.

(iii) H is termed an *n-parallel grammar*, *nL grammar* for short, if $K = \bigcup_{i=1}^n \bar{Y}_i^+$, for some $Y_i \subseteq \Sigma$, $1 \leqslant i \leqslant n$. H is *parallel* (H is an *L grammar*), if it is n-parallel for some $n \geqslant 1$.

(iv) H is termed an *n-sequential m-continuous grammar*, *nSmC grammar* for short, if $K = \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^* \cup \bigcup_{i=1}^m P_t \bar{Q}_i^+ R_i^*$, for some X_i, Y_i, Z_i, P_j, Q_j, $R_j \subseteq \Sigma$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$. H is an *SC grammar* if it is n-sequential m-continuous for some $n, m \geqslant 1$.

(v) H is termed an *n-sequential m-parallel grammar*, *nSmL grammar* for short, if $K = \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^* \cup \bigcup_{i=1}^m \bar{Q}_i^+$, for some X_i, Y_i, Z_i, $Q_j \subseteq \Sigma$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$. H is an *SL grammar* if it is n-sequential m-parallel for some $n, m \geqslant 1$. ▮

The following should be observed.

(1) In Ehrenfeucht *et al.* (1980) L grammars were called "simple continuous" (SC) grammars.

(2) We do not define n-continuous m-parallel grammars, since they are by definition $(n+m)$-continuous grammars.

The families of languages generated by nS, nC, nL, $nSmC$, $nSmL$ grammars, are denoted $\mathscr{L}(nS)$, $\mathscr{L}(nC)$, $\mathscr{L}(nL)$, $\mathscr{L}(nSmC)$, $\mathscr{L}(nSmL)$, respectively. Furthermore, $\mathscr{L}(S) = \bigcup_{n=1}^\infty \mathscr{L}(nS)$, $\mathscr{L}(C) = \bigcup_{n=1}^\infty \mathscr{L}(nC)$, $\mathscr{L}(L) = \bigcup_{n=1}^\infty \mathscr{L}(nL)$, $\mathscr{L}(SC) = \bigcup_{n=1}^\infty \bigcup_{m=1}^\infty \mathscr{L}(nSmC)$ and $\mathscr{L}(SL) = \bigcup_{n=1}^\infty \bigcup_{m=1}^\infty \mathscr{L}(nSmL)$.

If H is an S grammar (C grammar, SC grammar, SL grammar) and $X^* \bar{Y} Z^* \subseteq sel\ H$ (resp. $X^* \bar{Y}^+ Z^* \subseteq sel\ H$) we may informally refer to the elements of X and Z as (left and right, respectively) context-symbols.

If w is a word under rewriting, then all occurrences of symbols that will remain untouched are referred to as context.

If $v \in X^*\bar{Y}Z^*$ (resp. $v \in X^*\bar{Y}^+Z^*$) and $w = iden\ v = xyz$, with $x \in X^*$, $y \in Y$ (resp. $y \in Y^+$) and $z \in Z^*$, we call x and z a (permitting) context of y.

From the definitions it easily follows that $\mathscr{L}(CF) \subseteq \mathscr{L}(1S)$.

We end this section providing an example of a $2S$ grammar generating a non-$E0L$ language.

EXAMPLE II.1. We define the $2S$ grammar H by $H = (\{a, b, c, d\}, h, c, \{a, b\}, K)$, $K = K_1 \cup K_2$ with $K_1 = \{a, b, d\}^* \{\bar{a}, \bar{b}, \bar{c}\}\{c\}^*$ and $K_2 = \{c\}^*\{\bar{d}\}\{d\}^*$. h is defined by $h(c) = \{a, d\}$, $h(d) = \{c^2\}$, $h(a) = \{ab, ba\}$ and $h(b) = \{b^2\}$. Informally speaking, derivations in H proceed as follows.

The axiom c directly derives either a or d.

If a, then, since only K_1 can be applied, for every w such that $a \Rightarrow_H^* w$, $w = v_1 a$ or $w = v_2 b$ with $v_1 \in b^*$ and $v_2 \in b^*ab^*$ and only the last letter of w can be rewritten in a next derivation step.

Hence it is easy to see that $a \Rightarrow_H^* w$ if and only if $w \in L$, where $L = \{b^n ab^m : n \geqslant 0,\ m \geqslant 0\}$.

If $c \Rightarrow_H d$, then the next step yields c^2. Since in c^2 only the leftmost occurrence of c may be rewritten, we obtain $\{ac, dc\}$.

Inspecting ac and K yields that $ac \Rightarrow^* w_1 c$, with $w_1 \in L$ and $w_1 c \Rightarrow^* w_1 w_2$ with $w_2 \in L$ or $w_1 c \Rightarrow w_1 d$.

From the definition of K it follows that the non-terminal symbol d only can be rewritten if it occurs in a word u, with $alph\ u \subseteq \{c, d\}$. Since a and b never yield c or d or Λ it follows that no terminal words can be derived in H from words v with $\{a, d\} \subseteq alph\ v$. Hence $ac \Rightarrow^* w \in \Delta^*$ if and only if $w \in L \cdot L$.

Clearly dc yields da or dd.

From the above it follows that from da no terminal words can be derived in H. dd can only be rewritten according to K_2, yielding c^2d and to c^2d again only K_2 can be applied yielding c^2c^2.

Hence $c^2 \Rightarrow^* L \cdot L$ or $c^2 \Rightarrow dc \Rightarrow dd \Rightarrow c^2d \Rightarrow c^4$ are the only derivation sequences in H, which may yield terminal words.

It can now easily be seen that in general c^{2^k}, $k \geqslant 0$, either yields L^{2^k} or, as an intermediate step in a successful derivation, $c^{2^{k+1}}$.

Based on the above it follows that

$$L(H) = \{w \in L^{2^k} : k \geqslant 0\} = \{w \in \{a, b\}^* : \#_a w = 2^k,\ k \geqslant 0\}. \quad \blacksquare$$

III. The Generative Power of Sequential, Continuous and Parallel Rewriting

In this section we investigate the generative power of S grammars, C grammars and L grammars. In particular we investigate the mutual relationship between classes of languages generated by these classes of grammars, as well as we compare them with several known classes of languages.

We start by recalling several results from Ehrenfeucht *et al.* 1980).

THEOREM III.1. (i) $\mathscr{L}(1L) = \mathscr{L}(E0L)$.

(ii) $\mathscr{L}(L) = \mathscr{L}(ET0L)$.

THEOREM III.2. (i) $\mathscr{L}(1C) \setminus \mathscr{L}(ET0L) \neq \varnothing$.

(ii) $\mathscr{L}(1C) \supsetneq \mathscr{L}(E0L)$.

(iii) $\mathscr{L}(C) \supsetneq \mathscr{L}(ET0L)$.

We turn now to the investigation of $\mathscr{L}(S)$.

THEOREM III.3. (i) $\mathscr{L}(1S) \supseteq \mathscr{L}(CF)$.

(ii) $\mathscr{L}(2S) \supsetneq \mathscr{L}(CF)$.

Proof. This follows immediately from the definitions and Example II.1. ∎

We are not able to establish the precise relationship between $\mathscr{L}(1S)$ and $\mathscr{L}(CF)$.

However, we will investigate now a "natural" restriction on the form of a $1S$ grammar which guarantees that the language generated is context-free.

LEMMA III.1. *Let $H = (\Sigma, h, S, \Delta, X^*\bar{Y}\ Z^*)$ be a $1S$ grammar. Let $u \in$ sent H and $u \Rightarrow_H v$, such that $u = xyz$, with $x \in X^*$, $y \in Y$ and $z \in Z^*$, and $v = x\alpha z$, with $\alpha \in h(y)$. There exists no w, such that $v \Rightarrow_H^+ w$ if and only if one of the following conditions holds.*

(1) $v \in (\Sigma \setminus Y)^*$.

(2) $v \in \Sigma^*(X \setminus (Y \cup Z))(\Sigma \setminus Y)^*(Z \setminus (X \cup Y))\, \Sigma^*$.

(3) $\alpha \in \Sigma^*(\Sigma \setminus (X \cup Y \cup Z))\, \Sigma^*$.

(4) $\alpha \in \Sigma^*(Y \setminus (X \cup Z))\, \Sigma^*(Y \setminus (X \cup Z))\, \Sigma^*$.

(5) $\alpha \in \Sigma^*((Y \cup Z) \setminus X)\, \Sigma^*((X \cup Y) \setminus Z\, \Sigma^*$.

Proof. The "if" part of the statement is easily proved and hence left to the reader. The "only if" part is proved as follows.

(a) Suppose that we have a string s that cannot be rewritten. Since h is total, this means that there does not exist a word $t \in K$, such that *iden* $t = s$. We can distinguish four cases, described intuitively as follows.

(i) In s a symbol occurs that is not in $X \cup Y \cup Z$.

(ii) No symbol from Y occurs in s.

(iii) s contains two occurrences of symbols from Y, both of which are not context symbols.

(iv) If $s \in (X \cup Y \cup Z)^*$ and s contains at most one symbol from $Y \setminus (X \cup Z)$, then the context prescribes in a "contradictory way" which symbol must be rewritten.

Either (1) s contains a symbol from $(Z \cup Y) \setminus X$ and to the right of it an element from $(X \cup Y) \setminus Z$ occurs, or (2) s contains a symbol from $X \setminus (Y \cup Z)$ and to the right of it an element from $Z \setminus (X \cup Y)$ occurs and in between those two symbols no symbols from Y occur.

(*b*) Let u and v be as in the statement of the lemma. Hence $v = x\alpha z$, with $x \in X^*$, $\alpha \in h(y)$ and $z \in Z^*$, but v cannot be rewritten. Consequently, by (a), either $v \in (\Sigma \setminus Y)^* \cup \Sigma^*(X \setminus (Y \cup Z))\,(\Sigma \setminus Y)^*\,(Z \setminus (X \cup Y))\,\Sigma^*$ or α is of one of the following forms.

$$\alpha \in \Sigma^*(\Sigma \setminus (X \cup Y \cup Z))\,\Sigma^*,$$

$$\alpha \in \Sigma^*(Y \setminus (X \cup Z))\,\Sigma^*(Y \setminus (X \cup Z))\,\Sigma^*$$

or

$$\alpha \in \Sigma^*((Y \cup Z) \setminus X))\,\Sigma^*((X \cup Y) \setminus Z))\,\Sigma^*. \quad \blacksquare$$

LEMMA III.2. *For every* 1S *grammar H, there exists an equivalent* 1S *grammar $H' = (\Sigma, h, S, \Delta, K)$, with $K = X^*\bar{Y}Z^*$ such that $h(a) = \{a\}$ for all $a \in \Sigma \setminus Y$ and $Y \cap \Delta = \varnothing$.*

Proof. Let $H = (\Gamma, g, S', \Delta, K')$, with $K' = P^*\bar{Q}R^*$ be an arbitrary $1S$ grammar. We define the $1S$ grammar $H' = (\Sigma, h, S, \Delta, K)$ in the following way.

g' is the finite substitution from Γ^* into Γ^* defined by $g'(a) = g(a)$ if $a \in Q$ and $g'(a) = \{a\}$ if $a \notin Q$.

Obviously $(\Gamma, g', S', \Delta, K')$ generates $L(H)$.

$\Xi = \{\hat{a} : a \in \Delta \cap Q\}$ is an alphabet, disjoint with Γ. Let $\Sigma = \Gamma \cup \Xi$. The finite substitution φ from Γ^* into Σ^* is defined by $\varphi(a) = \{a\}$ if $a \notin \Delta \cap Q$ and $\varphi(a) = \{a, \hat{a}\}$ if $a \in \Delta \cap Q$.

Let h be the finite substitution from Γ^* into Γ^* defined by

$$h(\hat{a}) = \{a\} \cup \varphi g'(a) \qquad \text{for all } \hat{a} \in \Xi,$$

$$h(a) = \{a\} \qquad \text{for all } a \in \Delta \cup (\Gamma \setminus Q),$$

$$h(a) = \varphi g'(a) \qquad \text{for all } a \in Q \setminus \Delta.$$

Let $X = P \cup \{\hat{a} : a \in P\}$, $Z = R \cup \{\hat{a} : a \in R\}$ and $Y = (Q \setminus \Delta) \cup \Xi$. $S = S'$ if $S' \in \Gamma \setminus (Q \cap \Delta)$, $S = \hat{S}'$ if $S' \in Q \cap \Delta$,

Obviously H' satisfies all required restrictions.

The equivalence of H and H' can easily be proved and this is left to the reader. ∎

LEMMA III.3. *For every* 1S *grammar* H, *there exists an equivalent* 1S *grammar* $H' = (\Sigma, h, S, \Delta, K)$, *with* $K = X^* \bar{Y} Z^*$, *such that* $Y = \Sigma \setminus \Delta$ *and for all* $a \in \Delta$, $h(a) = \{a\}$.

Proof. Let $H = (\Gamma, g, S, \Delta, K)$ be a 1S grammar, with $K = P^* \bar{Q} R^*$, such that $g(a) = \{a\}$ for all $a \in \Gamma \setminus Q$ and $Q \cap \Delta = \varnothing$. Because of Lemma III.2 this assumption is not a restriction. Hence $Q \subseteq \Gamma \setminus \Delta$.

Let $\Sigma = Q \cup \Delta$ and let h be defined in the following way. For all $a \in \Sigma$ $h(a) = \{\alpha : \alpha \in g(a)$ and *alph* $\alpha \subseteq \Sigma\}$ if this is a nonempty set and $h(a) = \{a\}$ otherwise. Hence h is a finite substitution from Σ^* into Σ^*. Let $X = P \cap \Sigma$, $Z = R \cap \Sigma$ and $Y = Q$.

$H' = (\Sigma, h, S, \Delta, X^* \bar{Y} Z^*)$.

Clearly $L(H) = L(H')$ and hence the lemma holds. ∎

THEOREM III.4. *If* $H = (\Sigma, h, S, \Delta, X^* \bar{Y} Z^*)$ *is a* 1S *grammar, such that* $X \subseteq Y \cup Z$ *and* $Z \subseteq X \cup Y$, *then* $L(H) \in \mathscr{L}(CF)$.

Proof. By Lemma III.3 we may assume that H is such that $Y = \Sigma \setminus \Delta$, $h(a) = \{a\}$ for all $a \in \Delta$ and $a \in h(a)$ for all $a \in \Sigma$. We may assume also that the only production in which S occurs at the right-hand side is the identity production for S. Notice that the constructions used in the proofs of Lemmas III.2 and III.3 ensure that the condition $X \subseteq Y \cup Z$ and $Z \subseteq X \cup Y$ remains valid. This condition implies that

$$\Sigma^*(X \setminus (Y \cup Z))(\Sigma \setminus Y)^*(Z \setminus (X \cup Y))\, \Sigma^* = \varnothing.$$

The situation can be illustrated as in Fig. 1, where $\Sigma = (\Sigma \setminus \Delta) \cup \Delta = Y \cup \Delta$. The shadowed areas are empty and so

$$(X \cap Z) \setminus Y = X \cap \Delta = Z \cap \Delta. \tag{1}$$

By Lemma III.1 it is easily seen that we may assume the following. If $a \in \Sigma \setminus \Delta$ and α has one of the following three forms:

(i) $\alpha \in \Sigma^*(\Sigma \setminus (X \cup Y \cup Z))\, \Sigma^*$ and $\alpha \notin \Delta^+$,

(ii) $\alpha \in \Sigma^*(Y \setminus (X \cup Z))\, \Sigma^*(Y \setminus (X \cup Z)\, \Sigma^*$,

(iii) $\alpha \in \Sigma^*((Y \cup Z) \setminus X)\, \Sigma^*((X \cup Y) \setminus Z)\, \Sigma^*$,

then $\alpha \notin h(a)$.

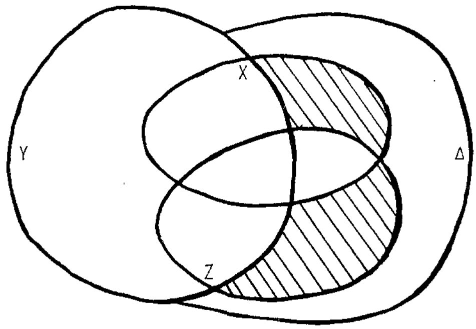

FIGURE 1

Hence a word $w \in sent\ H$ cannot derive any other word (but itself) if and only if either $w \in (\Sigma \setminus Y)^* = \Delta^*$ or in the last step in the derivation of w from S a production with right-hand side $\alpha \in \Delta^* \cap \Sigma^*(\Sigma \setminus (X \cup Y \cup Z))\, \Sigma^*$ has been applied.

Let $a \in \Sigma$. We now divide $h(a)$ by

$$h_1(a) = \{\alpha \in h(a) : alph\ \alpha \subseteq X \cup Y \cup Z\}$$

and

$$h_2(a) = h(a) \setminus h_1(a).$$

By the above it follows that $h_2(a) \subseteq \Delta^+$.

We can distinguish two types of successful derivations in H, those in which no production from h_2 is used and those in which productions from h_2 are used. We consider now derivations of the first type. By (1) we have that if $\alpha \in h_1(a)$ for some $a \in Y$ and $\alpha \in \Delta^+$ then $\alpha \in (X \cap Z)^+$. (Stronger: *alph* $\alpha \cap \Delta \subseteq X \cap Z$).

We can picture this situation as in Fig. 2.

If we now define the $E0S$ system G' by $G' = (\Sigma, h_1, S, \Delta)$, then it is rather easy to prove that $L(G') = \{w \in L(H) :$ no productions from h_2 are used in $S \Rightarrow_H^* w\}$.

Hence this part of $L(H)$ is a context-free language.

On the other hand if a production from h_2 was used in a successful derivation, then, as noticed before, this has been the only application of such a production in this derivation and it has been used in the last derivation step. Hence we can picture this situation as in Fig. 3.

We will construct an $E0S$ system G of which the already mentioned $E0S$ system G' is a subsystem. G will extend G' in the sense that exactly all derivations in G' are derivations in G and all derivations in G which are not derivations in G' have an analogon in the derivations in H in which h_2 has

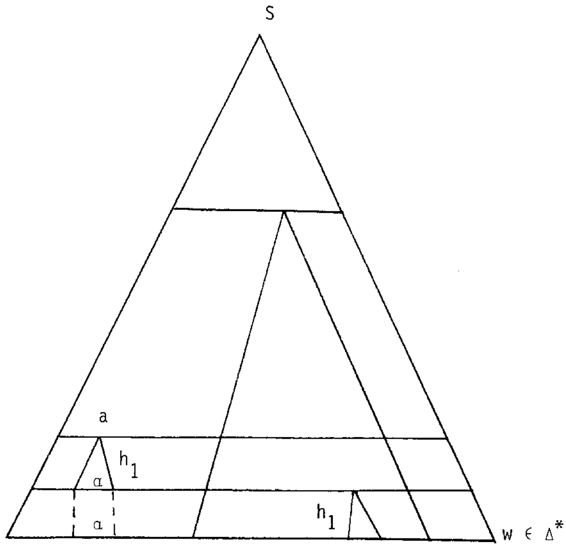

FIGURE 2

been used. It is rather clear (see Fig. 3) that if we consider a derivation in G which uses h_2, then in each sentential form of it we can distinguish one occurrence of a symbol which is an ancestor of the element of $\Sigma \setminus \Delta$ to which this production will be applied. This can easily be simulated in G. We

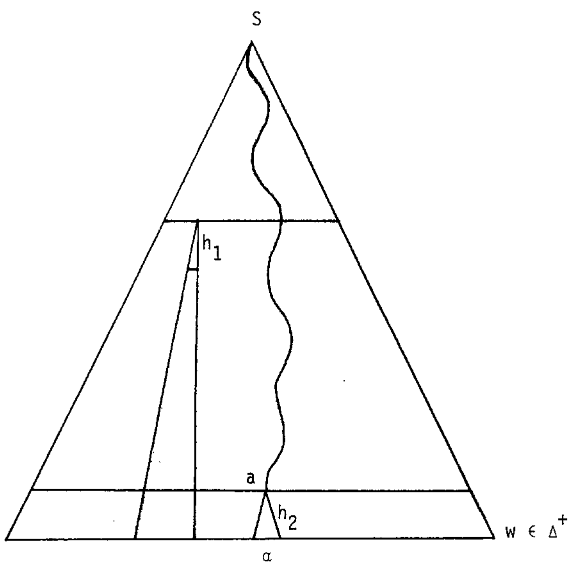

FIGURE 3

still have to take care of the fact that if such an ancestor y occurs in a string, then it should not produce a terminal string (by h_2) too soon.

If there are any other nonterminals in the string under rewriting, which require, according to H a context, then y must produce this context before a terminal string is derived. To this aim we introduce new symbols, which will remember their "context duties." These symbols will act as the ancestors of the symbol to which h_2 may be applied.

To all original symbols only h_1 is applicable and since they will eventually yield terminal words in $(X \cap Z)^*$ and since the ancestors of the symbol to which h_2 may be applied are subject to "context conditions" again an inductive argument is enough to prove that $L(G) = L(H)$.

Formally we define G as follows.

Let $[Y] = \{[a] : a \in Y\}$, $[Y, X] = \{[a, X] : a \in Y \setminus X\}$, $[Y, Z] = \{[a, Z] : a \in Y \setminus Z\}$ and $[Y, X, Z] = \{[a, X, Z] : a \in Y \setminus (X \cup Z)\}$ be new mutually disjoint alphabets.

Let $\hat{S}$ be a new symbol. $\Gamma = \Sigma \cup \{\hat{S}\} \cup [Y] \cup [Y, X] \cup [Y, Z] \cup [Y, X, Z]$. The finite substitution g from Γ^* into Γ^* is defined in the following way.

$g(\hat{S}) = \{S\} \cup$
$\{\alpha_1[a]\alpha_2 : \alpha_1 a \alpha_2 \in h_1(S), a \in X \cap Y \cap Z\} \cup$
$\{\alpha_1 [a]\alpha_2 : \alpha_1 a \alpha_2 \in h_1(S), a \in Y, \mathit{alph}\ \alpha_1\alpha_2 \cap Y = \varnothing\} \cup$
$\{\alpha_1[a, X]\alpha_2 : \alpha_1 a \alpha_2 \in h_1(S), a \in Y \setminus X, \mathit{alph}\ \alpha_2 \cap Y \neq \varnothing\} \cup$
$\{\alpha_1[a, Z]\alpha_2 : \alpha_1 a \alpha_2 \in h_1(S), a \in Y \setminus Z, \mathit{alph}\ \alpha_1 \cap Y \neq \varnothing\} \cup$
$\{\alpha_1[a, X, Z]\alpha_2 : \alpha_1 a \alpha_2 \in h_1(S), a \in Y \setminus (X \cup Z), \mathit{alph}\ \alpha_1 \cap Y \neq \varnothing$ and $\mathit{alph}\ \alpha_2 \cap Y \neq \varnothing\}$.
$g(S) = h(S)$.

For all $a \in \Sigma \setminus \{S\}$, $g(a) = h_1(a)$.
For all $[a] \in [Y]$,

$g([a]) = h_2(a) \cup$
$\{\alpha_1[b]\alpha_2 : \alpha_1 b \alpha_2 \in h_1(a), b \in X \cap Y \cap Z\} \cup$
$\{\alpha_1[b]\alpha_2 : \alpha_1 b \alpha_2 \in h_1(a), b \in Y, \mathit{alph}\ \alpha_1\alpha_2 \cap Y = \varnothing\} \cup$
$\{\alpha_1[b, X]\alpha_2 : \alpha_1 b \alpha_2 \in h_1(a), b \in (Y \cap Z) \setminus X, \mathit{alph}\ \alpha_2 \cap Y \neq \varnothing$ or $b \in Y \setminus X$, $\mathit{alph}\ \alpha_2 \cap Y \neq \varnothing$, $\mathit{alph}\ \alpha_1 \cap Y = \varnothing\} \cup$
$\{\alpha_1[b, Z]\alpha_2 : \alpha_1 b \alpha_2 \in h_1(a), b \in (X \cap Y) \setminus Z, \mathit{alph}\ \alpha_1 \cap Y \neq \varnothing$ or $b \in Y \setminus Z$, $\mathit{alph}\ \alpha_1 \cap Y \neq \varnothing$, $\mathit{alph}\ \alpha_2 \cap Y = \varnothing\} \cup$
$\{\alpha_1[b, X, Z]\alpha_2 : \alpha_1 b \alpha_2 \in h_1(a), b \in Y \setminus (X \cup Z), \mathit{alph}\ \alpha_1 \cap Y \neq \varnothing$ and $\mathit{alph}\ \alpha_2 \cap Y \neq \varnothing\}$.

For all $[a, X] \in [Y, X]$,

$$\begin{aligned}
g([a, X]) = {} & \{\alpha_1[b]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in X \cap Y \cap Z\} \cup \\
& \{\alpha_1[b]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in X \cap Y, \mathit{alph}\ \alpha_1 \cap Y = \varnothing\} \cup \\
& \{\alpha_1[b, Z]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in (X \cap Y) \setminus Z, \mathit{alph}\ \alpha_1 \cap Y \neq \varnothing\} \cup \\
& \{\alpha_1[b, X]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in (Y \cap Z) \setminus X\} \cup \\
& \{\alpha_1[b, X]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in Y \setminus (X \cup Z), \mathit{alph}\ \alpha_1 \cap Y = \varnothing\} \cup \\
& \{\alpha_1[b, X, Z]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in Y \setminus (X \cup Z), \mathit{alph}\ \alpha_1 \cap Y \neq \varnothing\}.
\end{aligned}$$

For all $[a, Z] \in [Y, Z]$,

$$\begin{aligned}
g([\mathrm{a}, Z]) = {} & \{\alpha_1[b]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in X \cap Y \cap Z\} \cup \\
& \{\alpha_1[b]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in Y \cap Z, \mathit{alph}\ \alpha_2 \cap Y = \varnothing\} \cup \\
& \{\alpha_1[b, X]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in (Y \cap Z) \setminus X, \mathit{alph}\ \alpha_2 \cap Y \neq \varnothing\} \cup \\
& \{\alpha_1[b, Z]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in (Y \cap Z) \setminus X\} \cup \\
& \{\alpha_1[b, Z]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in Y \setminus (X \cup Z), \mathit{alph}\ \alpha_2 \cap Y = \varnothing\} \cup \\
& \{\alpha_1[b, X, Z]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in Y \setminus (X \cup Z), \mathit{alph}\ \alpha_2 \cap Y \neq \varnothing\}.
\end{aligned}$$

For all $[a, X, Z] \in [Y, X, Z]$,

$$\begin{aligned}
g([a, X, Z]) = {} & \{\alpha_1[b]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in X \cap Y \cap Z\} \cup \\
& \{\alpha_1[b, X]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in (Y \cap Z) \setminus X\} \cup \\
& \{\alpha_1[b, Z]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in (X \cap Y) \setminus Z\} \cup \\
& \{\alpha_1[b, X, Z]\alpha_2 : \alpha_1 b\alpha_2 \in h_1(a), b \in Y \setminus (X \cup Z)\}.
\end{aligned}$$

Let G be defined by $G = (\Gamma, g, \hat{S}, \Delta)$. Then G is an $E0S$ grammar. Taking into account our intuitive description of how G simulates H and the formal definition of H, the reader should be able to produce a (rather tedious) formal proof of the equality $L(G) = L(H)$. From this equality it follows that $L(H) \in \mathscr{L}(CF)$. ∎

COROLLARY III.1. *A language is context-free if and only if it is generated by a* 1S *grammar H, such that sel $H = X^*\bar{Y}Z^*$ and $X \subseteq Y \cup Z$ and $Z \subseteq X \cup Y$.*

Proof. This statement follows immediately from the above theorem and the fact that the language generated by an $E0S$ system (Σ, h, S, Δ) is also generated by the 1S grammar $(\Sigma, h, S, \Delta, \Sigma^*\bar{\Sigma}\Sigma^*)$. ∎

In particular the above result implies that if all terminal symbols are both left and right context-symbols then the language generated is context-free.

COROLLARY III.2. *If* $H=(\Sigma, h, S, \Delta, X^*\bar{Y}Z^*)$ *is a* $1S$ *grammar, such that* $\Delta \subseteq X$ *and* $\Delta \subseteq Z$, *then* $L(H) \in \mathscr{L}(CF)$.

Proof. Combining Lemmas III.2 and III.3 we construct a $1S$ grammar $G=(\Gamma, g, Z, \Delta, P^*\bar{Q}R^*)$ equivalent with H, such that $\Delta \subseteq P$, $\Delta \subseteq R$ and $Q=\Gamma \diagdown \Delta$.

Then $P \subseteq \Gamma = \Gamma \diagdown \Delta \cup \Delta \subseteq Q \cup R$ and analogously $R \subseteq P \cup Q$.

Hence Theorem III.4 is applicable and the corollary holds. ∎

It is instructive to see that the above restriction on the "distribution" of terminal symbols does not imply context-freeness when one considers $2S$ rather than $1S$ grammars.

EXAMPLE III.1. Let $H=(\Sigma, h, S, \Delta, K)$ be a $2S$ grammar, defined in the following way.

$$\Sigma = \{S, A, A_1, B, B_1, C, C_1, D, D_1, a, b, c\},$$
$$\Delta = \{a, b, c\},$$

$$h(S)=\{AC\}, \qquad h(A)=\{aA_1b, aB\}, \qquad h(A_1)=\{A\},$$
$$h(B)=\{B_1\}, \qquad h(B_1)=\{b\},$$
$$h(C)=\{C_1c, D\}, \qquad h(C_1)=\{C\},$$
$$h(D)=\{D_1\}, \qquad h(D_1)=\{c\},$$
$$h(a)=\{a\}, \qquad h(b)=\{b\} \qquad \text{and} \qquad h(c)=\{c\}.$$

$K=K_1 \cup K_2$, where

$$K_1=\{a, b, c, A, B, B_1\}^* \, \{\bar{A}, \bar{C}_1, \bar{D}, \bar{B}, \bar{B}_1, \bar{S}\} \, \{a, b, c, C\}^*$$

and

$$K_2=\{a, b, c, A_1, B, B_1\}^* \, \{\bar{A}_1, \bar{C}, \bar{D}_1, \bar{B}, \bar{B}_1\} \, \{a, b, c, C_1\}^*.$$

From the construction of K it follows that Δ is a permitting context. That $L(H)=\{a^nb^nc^m : m \geqslant n \geqslant 1\}$ can be seen as follows.

Let $a^nAb^nCc^n \in \textit{sent } H$ for some $n \geqslant 0$ (notice that AC is the only word that can directly be derived from S).

Then $a^nAb^nCc^n \Rightarrow a^{n+1}A_1b^{n+1}Cc^n$ and $a^nAb^nCc^n \Rightarrow a^{n+1}Bb^nCc^n$ are all possible direct derivations from this word.

Either $a^{n+1}A_1b^{n+1}Cc^n \Rightarrow a^{n+1}A_1b^{n+1}C_1c^{n+1} \Rightarrow a^{n+1}Ab^{n+1}C_1c^{n+1} \Rightarrow a^{n+1}Ab^{n+1}Cc^{n+1}$ or $a^{n+1}A_1b^{n+1}Cc^n \Rightarrow a^{n+1}A_1b^{n+1}Dc^n$ whose last word cannot be rewritten anymore. In case the derived word is $a^{n+1}Bb^nCc^n$ we make the following observations.

In any $w \in \textit{sent } H$ at most two non-terminal symbols occur.

If B or $B_1 \in alph\ w$, they occur as the leftmost non-terminal.

If w contains any other non-terminal Q, then this $Q \in \{C, C_1, D, D_1\}$.

From the construction it follows that if $Q \in \{C, C_1\}$, then B and B_1 can be rewritten into B_1 and b, respectively, at any moment in the derivation and if $Q \in \{D, D_1\}$, then first D or D_1 is rewritten into D_1 or c, respectively, and after this B or B_1 can terminate.

Hence, either

$$a^{n+1}Bb^nCc^n \Rightarrow^* a^{n+1}b^{n+1}Qc^{n+k}, \qquad Q \in \{C, C_1\}, k \geqslant 0,$$

or

$$a^{n+1}Bb^nCc^n \Rightarrow^* a^{n+1}Bb^nDc^{n+k} \Rightarrow a^{n+1}Bb^nD_1c^{n+k}$$
$$\Rightarrow a^{n+1}Bb^nc^{n+k+1} \Rightarrow a^{n+1}B_1b^nc^{n+k+1} \Rightarrow a^{n+1}b^{n+1}c^{n+k+1}$$

or

$$a^{n+1}Bb^nCc^n \Rightarrow^* a^{n+1}B_1b^nDc^{n+k} \Rightarrow a^{n+1}B_1b^nD_1c^{n+k}$$
$$\Rightarrow a^{n+1}B_1b^nc^{n+k+1} \Rightarrow a^{n+1}b^{n+1}c^{n+k+1}.$$

It is easy to see that $a^{n+1}b^{n+1}Qc^{n+k} \Rightarrow^* a^{n+1}b^{n+1}c^{n+k+l}$, $Q \in \{C, C_1\}$, $k \geqslant 0$, $l \geqslant 1$. Hence it follows by an inductive argument, that $L(H) = \{a^nb^nc^m : m \geqslant n \geqslant 1\}$. ∎

We will investigate now "normal forms" for L, C and S grammars.

The following two theorems are results from Ehrenfeucht *et al.* (1980).

THEOREM III.5. $\mathscr{L}(L) = \mathscr{L}(2L)$.

THEOREM III.6. $\mathscr{L}(C) = \mathscr{L}(2C) = \mathscr{L}(1C1L)$.

THEOREM III.7. $\mathscr{L}(S) = \mathscr{L}(3S) = \mathscr{L}(1S1L)$.

Proof. By definition we have $\mathscr{L}(3S) \subseteq \mathscr{L}(S)$. The whole statement of the theorem follows, when we have shown as well that

(i) $\mathscr{L}(nS) \subseteq \mathscr{L}(1S1L)$ for all $n \geqslant 1$ and

(ii) $\mathscr{L}(1S1L) \subseteq \mathscr{L}(3S)$.

(i) Let $H = (\Sigma, h, S, \Delta, K)$ be an nS grammar; $K = \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^*$. For $i = 1,\dots, n$, $\hat{X}_{\langle i\rangle} = \{\hat{a}_{\langle i\rangle} : a \in X_i\}$, $\dot{Y}_{\langle i\rangle} = \{\dot{a}_{\langle i\rangle} : a \in Y_i\}$, $\check{Z}_{\langle i\rangle} = \{\check{a}_{\langle i\rangle} : a \in Z_i\}$ and $\sum_{\langle i\rangle} = \{a_{\langle i\rangle} : a \in \Sigma\}$ are pairwise disjoint new alphabets. $\Gamma = \bigcup_{i=1}^n (\hat{X}_{\langle i\rangle} \cup \dot{Y}_{\langle i\rangle} \cup \check{Z}_{\langle i\rangle} \cup \Sigma_{\langle i\rangle}) \cup \Delta$.

The finite substitution g, from Γ^* into Γ^*, is defined as follows.

For $i = 2,\dots, n$, if $a \in X_i$, then $\hat{a}_{\langle i\rangle} \in g(a_{\langle i-1\rangle})$, $\hat{a}_{\langle i\rangle} \in g(\hat{a}_{\langle i-1\rangle})$, $\hat{a}_{\langle i\rangle} \in g(\check{a}_{\langle i-1\rangle})$, (if $a \in X_{i-1}$, Z_{i-1} resp. for the last two cases); if $a \in Y_i$, then $\dot{a}_{\langle i\rangle} \in g(a_{\langle i-1\rangle})$, $\dot{a}_{\langle i\rangle} \in g(\hat{a}_{\langle i-1\rangle})$, $\dot{a}_{\langle i\rangle} \in g(\check{a}_{\langle i-1\rangle})$, (if $a \in X_{i-1}$, Z_{i-1}, resp. for the last two cases); if $a \in Z_i$, then $\check{a}_{\langle i\rangle} \in g(a_{\langle i-1\rangle})$, $\check{a}_{\langle i\rangle} \in g(\hat{a}_{\langle i-1\rangle})$ and $\check{a}_{\langle i\rangle} \in g(\check{a}_{\langle i-1\rangle})$, (if $a \in X_{i-1}$, Z_{i-1} resp. for the last two cases); if $a \in \Delta$, then $a \in g(a_{\langle i\rangle})$ and $a \in g(a)$, if $a \in Y_i$ and $b_1 \cdots b_k \in h(a)$ for some $b_j \in \Sigma$, $1 \leqslant j \leqslant k$, then $b_{1\langle i\rangle} \cdots b_{k\langle i\rangle} \in g(\dot{a}_{\langle i\rangle})$; if $a \in Y_i$ and $\Lambda \in h(a)$, then $\Lambda \in g(\dot{a}_{\langle i\rangle})$.

To obtain the case $i = 1$, replace in the above i by 1 and $i - 1$ by n. Let $K' = K_1 \cup K_2$, where $K_1 = (\bigcup_{i=1}^{n} \bar{\Sigma}_{\langle i\rangle} \cup \bar{X}_{\langle i\rangle} \cup \check{Z}_{\langle i\rangle})^+$ and

$$K_2 = \left(\bigcup_{i=1}^{n} \hat{X}_{\langle i\rangle}\right)^* \left(\bigcup_{i=1}^{n} \bar{Y}_{\langle i\rangle}\right) \left(\bigcup_{i=1}^{n} \check{Z}_{\langle i\rangle}\right)^*$$

$H' = (\Gamma, g, S_{\langle 1\rangle}, \Delta, K')$.

Obviously H' is a $1S1L$ grammar.

That $L(H') = L(H)$ can be seen, when the following facts are taken into consideration.

K_1 acts as a rotator of indexed letters. As long as no dotted or non-indexed symbols are introduced K_1 may be used. In each such rewriting all indices $i \leqslant n - 1$ are increased by 1 and the index n changes into 1. As soon as a non-indexed symbol is introduced the derivation has come to an end, since K' contains no words with such letters. When a dotted symbol is introduced K_2 must be used in the next derivation step, which implies that the word under rewriting must be of the form

$$xyz \text{ with } x \in \left(\bigcup_{i=1}^{n} \hat{X}_{\langle i\rangle}\right)^*, y \in \bigcup_{i=1}^{n} \dot{Y}_{\langle i\rangle}, z \in \left(\bigcup_{i=1}^{n} \check{Z}_{\langle i\rangle}\right)^*.$$

From the first part of this explanation it follows that in xyz only one index, say i_0, occurs and hence $xyz \in \hat{X}^*_{\langle i_0\rangle} \dot{Y}_{\langle i_0\rangle} \check{Z}^*_{\langle i_0\rangle}$.

Now $y = \dot{b}_{\langle i_0\rangle}$ will be rewritten either in $b_{1\langle i_0\rangle} \cdots b_{k\langle i_0\rangle}$ if $b_1 \cdots b_k \in h(b)$ or in Λ if $\Lambda \in h(b)$.

After this step (notice that in the resulting word still only the index i_0 occurs) K_1 has to be applied again. Using this informal exposition a formal proof of $L(H) = L(H')$ can easily be formulated.

(ii) Let $H = (\Sigma, h, S, \Delta, K)$ be a $1S1L$ grammar, $K = X^* \bar{Y}_1 Z^* \cup \bar{Y}_2^+$. Let Σ_1 and Σ_2 be new, mutually disjoint alphabets.

$$\Sigma_1 = \{\hat{a} : a \in Y_1\} \text{ and } \Sigma_2 = \{\check{a} : a \in Y_2\}.$$

$$\Gamma = \Sigma_1 \cup \Sigma_2 \cup \Sigma.$$

We define the finite substitution g from Γ^* into Γ^* as follows.

For all $a \in Y_1, \hat{a} \in g(a)$ and $g(\hat{a}) = h(a)$; for all $a \in Y_2, \check{a} \in g(a)$ and $g(\check{a}) = h(a)$; for all $a \in \Sigma \setminus (Y_1 \cup Y_2)$, $a \in g(a)$.

The selector K' is defined by

$$K' = X^*(\bar{Y}_1 \cup \bar{\Sigma}_1)Z^* \cup \Sigma_2^* \bar{\Sigma}\Sigma^* \cup \Sigma^* \bar{\Sigma}_2 \Sigma_2^* .$$

$H' = (\Gamma, g, S, \Delta, K')$ is a $3S$ grammar, generating $L(H)$.

(iii) Hence we have $\mathscr{L}(3\mathrm{S}) \subseteq \mathscr{L}(\mathrm{S}) \subseteq \mathscr{L}(1\mathrm{S}1\mathrm{L}) \subseteq \mathscr{L}(3\mathrm{S})$ and the theorem holds. ∎

Next we consider the relationship between $\mathscr{L}(C)$ and $\mathscr{L}(S)$.

LEMMA III.4. $\mathscr{L}(C) \subseteq \mathscr{L}(S)$.

Proof. Let $H = (\Sigma, h, S, \Delta, K)$ be an arbitrary C grammar. We may assume that $K = X^* \bar{Y}_1^+ Z^* \cup \bar{Y}_2^+$, for some $X, Y_1, Z, Y_2 \subseteq \Sigma$.

We define the following mutually disjoint, new alphabets.

$$\hat{X} = \{\hat{a} : a \in X\},\ Y_{\langle 1\rangle} = \{a_{\langle 1\rangle} : a \in Y_1\},\ \check{Z} = \{\check{a} : a \in Z\},$$

$$Y_{\langle 2\rangle} = \{a_{\langle 2\rangle} : a \in Y_2\}.$$

Let $\Gamma = \Sigma \cup \hat{X} \cup Y_{\langle 1\rangle} \cup \check{Z} \cup Y_{\langle 2\rangle}$. The finite substitution g from Γ^* into Γ^* is defined by

$$\begin{array}{lll}
\hat{a} \in g(a) & \text{if } a \in X, & \\
a_{\langle i\rangle} \in g(a) & \text{if } a \in Y_i, & \text{for } i \in \{1, 2\}, \\
\check{a} \in g(a) & \text{if } a \in Z, & \\
g(a) = \{a\} & \text{if } a \notin X \cup Y_1 \cup Z \cup Y_2, & \\
g(\hat{a}) = \{a\} & \text{for all } a \in X, & \\
g(\check{a}) = \{a\} & \text{for all } a \in Z, & \\
g(a_{\langle i\rangle}) = h(a) & \text{for all } a \in Y_i, \text{ for } i \in \{1, 2\}, &
\end{array}$$

Let $K' = (\Gamma \setminus \Sigma)^* \bar{\Sigma} \Sigma^* \cup \Sigma^* \bar{\hat{X}}(\hat{X} \cup Y_{\langle 1\rangle} \cup \check{Z})^* \cup \Sigma^* \bar{Y}_{\langle 1\rangle}(Y_{\langle 1\rangle} \cup \check{Z})^* \cup \Sigma^* \bar{\check{Z}} \check{Z}^* \cup \Sigma^* \bar{Y}_{\langle 2\rangle} Y_{\langle 2\rangle}^*$.

$$H' = (\Gamma, g, S, \Delta, K').$$

From the construction it is clear that H' is a $5S$ grammar and that $L(H') = L(H)$.

Hence the lemma holds. ∎

LEMMA III.5. $\mathscr{L}(S) \subseteq \mathscr{L}(C)$.

Proof. Let $H = (\Sigma, h, S, \Delta, K)$ be a $1S1L$ grammar, $K = X^*\bar{Y}_1Z^* \cup \bar{Y}_2^+$. Let $\hat{Y}_1 = \{\hat{a} : a \in Y_1\}$ and $\check{Y}_2 = \{\check{a} : a \in Y_2\}$ be new alphabets, mutually disjoint.

Let $\Gamma = \Sigma \cup \hat{Y}_1 \cup \check{Y}_2 \cup \{¢\}$, where $¢ \notin \Sigma \cup \hat{Y}_1 \cup \check{Y}_2$. We define g, a finite substitution from Γ^* into Γ^* as follows.

If $a \in Y_1$, then $¢\hat{a} \in g(a)$ and $g(\hat{a}) = h(a)$, if $a \in Y_2$, then $\check{a} \in g(a)$ and $g(\check{a}) = h(a)$, for all $a \in \Sigma \diagdown (Y_1 \cup Y_2)$, $g(a) = \{a\}$, $g(¢) = \{\Lambda\}$.

$K' = K_1 \cup K_2 \cup K_3$, where $K_1 = (X \cup ¢)^*(\bar{Y}_1 \cup \bar{\hat{Y}}_1)^+Z^*$, $K_2 = \Sigma^*\bar{¢}_1^+\Sigma^*$,

$K_3 = (\bar{Y}_2 \cup \bar{\check{Y}}_2)^+$,

$H' = (\Gamma, g, S, \Delta, K')$.

Notice that in case K_1 is used in a successful derivation, according to H', then only one symbol (and not a string of symbols) of Y_1 or $\hat{Y}_1$ is rewritten. Would there be more symbols from Y_1 rewritten, then there would be more than one ¢ (and symbols from $\hat{Y}_1$) in the derived string, which would be of the form $x¢\hat{y}_1¢\hat{y}_2\alpha z$ and which hence could not be rewritten anymore, in contradiction with the successfulness of the derivation. As a consequence in each intermediate word of the derivation at most one symbol of $\hat{Y}_1$ occurs. That K_1 and K_2 together simulate the sequential part of K is now clear. K_3 obviously simulates the parallel part of K. So we conclude that $L(H') = L(H)$. Combining Theorem III.7 with the above reasoning it follows that for every S grammar there exists an equivalent C grammar. ∎

THEOREM III.8. $\mathscr{L}(\mathrm{S}) = \mathscr{L}(\mathrm{C})$.

We can add to Lemma III.4. the following result, which gives more insight in the relationship between the families of languages generated by $1C$ and $2S$ grammars.

THEOREM III.9. $\mathscr{L}(1\mathrm{C}) \subseteq \mathscr{L}(2\mathrm{S})$.

Proof. Let $H = (\Sigma, h, S, \Delta, K)$ with $K = X^*\bar{Y}^+Z^*$ be a $1C$ grammar. Let $H = (\Gamma, g, S, \Delta, K')$ be the $2S$ grammar, defined as follows. $\Gamma = \Sigma \cup \hat{X} \cup \dot{Y}$, where $\hat{X} = \{\hat{a} : a \in X\}$ and $\dot{Y} = \{\dot{a} : a \in Y\}$ are mutually disjoint alphabets, also disjoint with Σ. g is the finite substitution, from Γ^* into Γ^*, defined by

$$\begin{array}{lll} g(a) = \{a\} & \text{if} \quad a \in \Sigma \diagdown (X \cup Y), & \\ \hat{a} \in g(a) & \text{and} \quad g(\hat{a}) = \{a\} & \text{if} \quad a \in X, \\ \dot{a} \in g(a) & \text{and} \quad g(\dot{a}) = h(a) & \text{if} \quad a \in Y. \end{array}$$

Let $K' = K_1 \cup K_2$, where $K_1 = (\hat{X} \cup \dot{Y})^*(\bar{\Sigma} \cup \bar{\hat{X}})\Sigma^*$ and $K_2 = (\hat{X} \cup \Sigma)^*$

$\bar{Y}(\dot{Y} \cup Z)^*$. The inclusion $L(H) \subseteq L(H')$ can easily be proved. That $L(H') \subseteq L(H)$ can be seen as follows.

Let $w \in \Sigma^*$ be an arbitrary sentential term of H' that can be further rewritten. Obviously any rewriting of w in H' starts according to K_1.

Then w yields a word $w' \in (\hat{X} \cup \dot{Y})\Sigma^*$. This implies that $w \in (X \cup Y)\Sigma^*$. Assume that w' is rewritten (several times) using K_1 and yielding w'' in $(\hat{X} \cup \dot{Y})^*\Sigma^*$. If K_2 is applied to w'', then $w'' \in \hat{X}^*\dot{Y}\dot{Y}^*Z^*$. Hence $w \in X^*\bar{Y}^+Z^*$.

Let $w'' = \hat{x}\dot{y}\dot{u}z$, with $\hat{x} \in \hat{X}^*$, $\dot{y} \in \dot{Y}$, $\dot{u} \in \dot{Y}^*$ and $z \in Z^*$ then $w'' \Rightarrow_{H'} \hat{x}\gamma\dot{u}z$, where $\gamma \in g(\dot{y}) = h(y)$. If $\gamma \neq \Lambda$ then $\hat{x}\gamma\dot{u}z$ can be rewritten according to K_2 only. If $\gamma = \Lambda$ then also K_1 may be applied. We have to consider two cases.

(1) Only elements from Z can get rewritten: $\hat{x}\dot{u}az' \Rightarrow_{H'} \hat{x}\dot{u}bz'$, where $b \in \{a, \hat{a}, \dot{a}\}$ and $z = az'$; hence $b = \dot{a}$. It is rather straightforward to see now that the sequence $w \Rightarrow^*_{H'} \hat{x}\dot{y}\dot{u}az' \Rightarrow_{H'} \hat{x}\dot{u}az' \Rightarrow_{H'} \hat{x}\dot{u}\dot{a}z'$ and the sequence $w \Rightarrow \hat{x}\dot{y}\dot{u}az' \Rightarrow \hat{x}\dot{y}\dot{u}\dot{a}z' \Rightarrow_{H'} \hat{x}\dot{u}\dot{a}z'$ lead to the same words in $L(H')$.

(2) $\hat{x} = \hat{x}'c$, with $\hat{c} \in \hat{X}$, and $\hat{x}\dot{u}z \Rightarrow_{H'} \hat{x}'c\dot{u}z$, which implies $\dot{u} = \Lambda$ yielding an analogon of a previous situation.

Based on the above remarks it is rather easy to construct a formal proof of the fact that $L(H') \subseteq L(H)$. Consequently the theorem holds. ∎

It is natural to consider selective substitution grammars which contain selectors of mixed categories. From Theorems III.5, III.6 and III.7 we have seen that augmenting $1L$, $1C$ or $1S$ grammars with one additional L selector allows one to generate $\mathscr{L}$(L), $\mathscr{L}$(C) and $\mathscr{L}$(S), respectively. The remaining situation of combining S and C selectors is considered now.

LEMMA III.6. *$\mathscr{L}(n\mathrm{S}m\mathrm{C}) \subseteq \mathscr{L}(\mathrm{S})$ for all integers n, m.*

Proof. Let H be an arbitrary $nSmC$ grammar. Since in Lemma III.4 it was proved that $\mathscr{L}(m\mathrm{C}) \subseteq \mathscr{L}(\mathrm{S})$ and since by definition $\mathscr{L}(n\mathrm{S}) \subseteq \mathscr{L}(\mathrm{S})$, we may assume that $n, m \geqslant 1$. So, $H = (\Sigma, h, S, \Delta, K)$ with $K = \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^* \cup \bigcup_{i=1}^m P_i^* \bar{Q}_i^+ R_i^*$.

Let $\Sigma_{\langle i \rangle} = \{a_{\langle i \rangle} : a \in Y_i\}$ for $i = 1,..., n$ and for $i = n+1,...\ n+m$, $\Sigma_{\langle i \rangle} = \{a_{\langle i \rangle} : a \in Q_{i-n}\}$ and $\Sigma_{l\langle i \rangle} = \{a_{l\langle i \rangle} : a \in P_{i-n}\}$, and $\Sigma_{r\langle i \rangle} = \{a_{r\langle i \rangle} : a \in R_{i-n}\}$, be mutually disjoint new alphabets. Let $\Gamma = \bigcup_{i=1}^{n+m} \Sigma_{\langle i \rangle} \cup \bigcup_{i=n+1}^{n+m} (\Sigma_{l\langle i \rangle} \cup \Sigma_{r\langle i \rangle}) \cup \Sigma$.

Let g be the finite substitution from Γ^* into Γ^* defined as follows.

For $1 \leqslant i \leqslant n$, if $a \in Y_i$, then $a_{\langle i \rangle} \in g(a)$ and $g(a_{\langle i \rangle}) = h(a)$; for $n+1 \leqslant i \leqslant n+m$,

$$\text{if } a \in Q_{i-n}, \text{ then } a_{\langle i \rangle} \in g(a) \text{ and } g(a_{\langle i \rangle}) = h(a),$$
$$\text{if } a \in P_{i-n}, \text{ then } a_{l\langle i \rangle} \in g(a) \text{ and } g(a_{l\langle i \rangle}) = \{a\},$$
$$\text{if } a \in R_{i-n}, \text{ then } a_{r\langle i \rangle} \in g(a) \text{ and } g(a_{r\langle i \rangle}) = \{a\},$$

and finally,

$$\text{if } a \in \Sigma \Big\backslash \left(\bigcup_{i=1}^{n} Y_i \cup \bigcup_{i=1}^{m} (P_i \cup Q_i \cup R_i) \right), \text{ then } g(a) = \{a\}.$$

$$K' = \bigcup_{i=1}^{n} K_i \cup \bigcup_{i=n+1}^{n+m+1} K_i,$$

where

$K_i = X_i^*(\bar{Y}_i \cup \bar{\Sigma}_{\langle i\rangle})Z_i^*$ for $1 \leqslant i \leqslant n$, $K_i = \Sigma^* \bar{\Sigma}_{l\langle i\rangle}(\Sigma_{l\langle i\rangle} \cup \Sigma_{\langle i\rangle} \cup \Sigma_{r\langle i\rangle})^* \cup \Sigma^* \bar{\Sigma}_{\langle i\rangle}(\Sigma_{\langle i\rangle} \cup \Sigma_{r\langle i\rangle})^* \cup \Sigma^* \bar{\Sigma}_{r\langle i\rangle}\Sigma_{r\langle i\rangle}^*$, for $n+1 \leqslant i \leqslant n+m$ and

$$K_{n+m+1} = \left(\bigcup_{i=n+1}^{n+m} \Sigma_{l\langle i\rangle} \cup \Sigma_{\langle i\rangle} \cup \Sigma_{r\langle i\rangle} \right)^* \bar{\Sigma}\Sigma^*.$$

Let $H' = (\Gamma, g, S, \Delta, K')$. To prove that $L(H') = L(H)$ we proceed as follows.

Let $w \in \Sigma^+$ be a sentential form of H'. Let $w \Rightarrow_{H'} v \Rightarrow_{H'} u \Rightarrow_{H'}^* w'$ be a derivation such that $w, w' \in \Sigma^*$ and all intermediate words contain symbols not from Σ. It is clear that, if $w = b_1 \cdots b_k$ for some $b_j \in \Sigma$, $1 \leqslant j \leqslant k$, then $v = b_1 \cdots b_{j-1} b b_{j+1} \cdots b_k$ for some $b \in \Gamma \setminus \Sigma$. For $v \Rightarrow u$ there are two possibilities for the choice of selector.

If K_i, with $1 \leqslant i \leqslant n$, is used, then $b_1 \cdots b_{j-1} \in X_i^*$, $b \in \Sigma_{\langle i\rangle}$ and $b_{j+1} \cdots b_k \in Z_i^*$ which implies $b_j \in Y_i$ and hence $w \in X_i^* Y_i Z_i^*$, and $u = b_1 \cdots b_{j-1} \gamma b_{j+1} \cdots b_k$ with $\gamma \in g(b) = h(b_j)$. Thus in this case

$$w \Rightarrow_H u = w'. \tag{1}$$

If K_i, with $n+1 \leqslant i \leqslant n+m+1$, is used then $v = bb_2 \cdots b_k$, if $k \geqslant 2$, and $v = b$ if $k = 1$, with $b \in \Sigma_{l\langle i_0\rangle} \cup \Sigma_{\langle i_0\rangle} \cup \Sigma_{r\langle i_0\rangle}$, for some $i_0 \in \{n+1,..., n+m\}$. If $k = 1$, then either $u = w$, if $b \in \Sigma_{l\langle i_0\rangle} \cup \Sigma_{r\langle i_0\rangle}$, or $u = \gamma$, if $b \in \Sigma_{\langle i_0\rangle}$, and $\gamma \in g(b) = h(b_1)$. If $k \geqslant 2$ then $u = bb'b_3 \cdots b_k$, with $b' \in \Sigma_{l\langle j_0\rangle} \cup \Sigma_{\langle j_0\rangle} \cup \Sigma_{r\langle j_0\rangle}$ or $b' \in \Sigma_{\langle j_1\rangle}$ for some $j_0 \in \{n+1,..., n+m\}$ and $j_1 \in \{1,..., n\}$.

Because all indexed alphabets are disjoint we must conclude that only the first possibility can lead to a successful derivation, and moreover $j_0 = i_0$. This reasoning holds also for the further rewriting of u and hence we have that all symbols of w are rewritten into "themselves with indices," giving w'.

Inspecting K' now, we see that w' has an ancestor of the form pqr with $p \in \Sigma_{l\langle i_0\rangle}^*$, $q \in \Sigma_{\langle i_0\rangle}^*$ and $r \in \Sigma_{r\langle i_0\rangle}^*$. Now it is easy to see that if $q = \Lambda$ then $w' = w$ and else $w' = b_1 \cdots b_l \gamma b_r \cdots b_k$, with $b_1 \cdots b_l \in P_{i_0}^*$, $\gamma \in g(q) = h(b_{l+1} \cdots b_{r-1})$, and $b_r \cdots b_k \in R_{i_0}$ and hence $w \Rightarrow_H w'$. (2)

Equations (1) and (2) imply that $L(H') \subseteq L(H)$.

The converse inclusion follows easily.

If we combine the equality $L(H')=L(H)$ with the fact that H' is an S grammar the theorem follows. ∎

THEOREM III.10. $\mathscr{L}(S)=\mathscr{L}(SC)=\mathscr{L}(C)$.

Proof. From Lemma III.6 we have that $\mathscr{L}(SC)\subseteq\mathscr{L}(S)$. By definition $\mathscr{L}(S)\subseteq\mathscr{L}(SC)$ and from Theorem III.8 we know that $\mathscr{L}(S)=\mathscr{L}(C)$. From this the theorem follows. ∎

IV. DETERMINISM

Given a selective substitution grammar $H=(\Sigma, h, S, \Delta, K)$ and a word $w\in\Sigma^+$, one of the following two situations can happen. (1) At most one $v\in K$ exists such that *iden* $v=w$, or (2) more than one $v\in K$ exist such that *iden* $v=w$.

In the former case, either w cannot be rewritten at all or otherwise the set of occurrences of letters in w that are rewritten is "deterministically determined." This does not mean that w, when rewritten, will have exactly one successor, because H may have several productions available for a symbol to be rewritten. Thus if H satisfies (1) for every $w\in\Sigma^+$, then we would like to talk about "global determinism," since it refers to a global property of the language K. Note that if H is a parallel grammar then obviously it is globally deterministic. If H is a "context-free" grammar (that is, $K=\Sigma^*\overline{\Sigma\setminus\Delta}\,\Sigma^*$) then it is not globally deterministic since one can choose an arbitrary occurrence of a non-terminal to rewrite. However, it is well known that each context-free language can be generated by a context-free grammar in a leftmost fashion; that is, using $K=\Delta^*\overline{\Sigma\setminus\Delta}\,\Sigma^*$. In this case we deal with a globally deterministic grammar.

In the case of a sequential, a continuous and a parallel grammar $H=(\Sigma, h, S, \Delta, K)$ the selector K is given in the union form $\bigcup_{i=1}^{n} K_i$. Clearly, in general, such a union form is not unique. Thus for our further considerations we will assume that a fixed representation $\{K_1,..., K_n\}$ of K is given, where $K=\bigcup_{i=1}^{n} K_i$.

Now we can distinguish another kind of determinism. Given a word $w\in\Sigma^+$, one of the following two situations can happen. (3) At most one $i\in\{1,..., n\}$ exists, such that K_i contains a word v, with *iden* $v=w$, or (4) more than one i with this property exists. In the former case either w cannot be rewritten at all or otherwise there exists exactly one component which may rewrite w. It does not mean that within this unique component only one v exists with *iden* $v=w$. If H satisfies (3) for every $w\in\Sigma^+$ then we would like to talk about "component determinism" of H.

These two notions of determinism are investigated in this section.

Formally we have the following definitions.

DEFINITION IV.1. Let $H=(\Sigma, h, S, \Delta, K)$ be an $E0S$ based s-grammar and let $\{K_1,\dots,K_n\}$ be a fixed representation of K; $K=\bigcup_{i=1}^{n} K_i$.

H is *globally deterministic* if for every $w \in \Sigma^+$ there exists at most one $v \in K$ such that *iden* $v=w$.

H is *component deterministic* if for every $w \in \Sigma^+$ there exists at most one $i \in \{1,\dots,n\}$ such that K_i contains a word v with *iden* $v=w$.

H is *selector deterministic* if H is both globally deterministic and component deterministic. ∎

We start by considering continuous and parallel grammars.

THEOREM IV.1. *A* $1C$ *grammar* H, *with sel* $H=X^*\bar{Y}^+Z^*$ *is globally deterministic if and only if* $X \cap Y = Y \cap Z = \varnothing$.

Proof. Let $H=(\Sigma, h, S, \Delta, X^*\bar{Y}^+Z^*)$. If $w \in \Sigma^*$ is such that there exist $v_1, v_2 \in$ *sel* H with $v_1 \neq v_2$ and *iden* $v_1=$ *iden* $v_2=w$, then $w=x_1y_1z_1=x_2y_2z_2$, with $x_1, x_2 \in X^*$, $y_1, y_2 \in Y^+$, $z_1, z_2 \in Z^*$ and $x_1y_1 \neq x_2y_2$ or $y_1z_1 \neq y_2z_2$. Hence either (1) $x_1 \neq x_2$ or (2) $z_1 \neq z_2$ or (3) both $x_1 \neq x_2$ and $z_1 \neq z_2$. Equivalently, either (1) $x_1=x_2\alpha$ or $x_2=x_1\alpha$, with $\alpha \in Y^+Z^*$, or (2) $z_1=\beta z_2$ or $z_2=\beta z_1$ with $\beta \in X^*\bar{Y}^+$, or (3) $x_1=x_2\alpha$ and $z_1=\beta z_2$ or $z_2=\beta z_1$, or $x_2=x_1\alpha$ and $z_1=\beta z_2$ or $z_2=\beta z_1$ with $\alpha \in \bar{Y}^+Z^*$ and $\beta \in X^*\bar{Y}^+$. It follows that *alph* $\alpha \cap X \cap Y \neq \varnothing$ and *alph* $\beta \cap Y \cap Z \neq \varnothing$. On the other hand if $X \cap Y \neq \varnothing$ or $Y \cap Z \neq \varnothing$ then, for all $w \in X^*(X \cap Y)^2Z^*$ or $w \in X\ (Y \cap Z)^2Z^*$, respectively, there exist more than one $v \in$ *sel* H such that *iden* $v=w$. This proves the theorem. ∎

COROLLARY IV.1. *A* $1C$ *grammar* H, *with sel* $H=X^*\bar{Y}^+Z^*$ *is selector deterministic if and only if* $X \cap Y = Y \cap Z = \varnothing$.

In Ehrenfeucht *et al.* (1980) a left- (or right-) continuous normal form was established for continuous grammars. (A C grammar H is said to be left- (or right-) continuous if *sel* $H=\bigcup_{i=1}^{n} \bar{Y}_i^+ Z_i^*$ (or $H=\bigcup_{i=1}^{n} X_i^* \bar{Y}_i^+$)).

If we combine their Theorem 1 and the construction from the proof of their Theorem 2 we obtain the following result.

THEOREM IV.2. *For every* C *grammar there exists an equivalent selector deterministic* C *grammar in left-(right-) continuous normal form.*

COROLLARY IV.2. *For every* L *grammar there exists an equivalent selector deterministic* L *grammar.*

We move now to consider sequential grammars. First of all, similarly as for $1C$ grammars, we can characterize global determinism in $1S$ grammars.

THEOREM IV.3. *A* $1S$ *grammar* H, *with sel* $H=X^*\bar{Y}Z^*$, *is globally deterministic if and only if* $X \cap Y = \varnothing$ *or* $Y \cap Z = \varnothing$.

Proof. Let H be defined by $H=(\Sigma, h, S, \Delta, X^*\bar{Y}Z^*)$. Let $w \in \Sigma^*$ and let $v_1, v_2 \in K$ be such that $v_1 \neq v_2$ and *iden* $v_1 =$ *iden* $v_2 = w$. Hence there exist $x_1, x_2 \in X^*$, $y_1, y_2 \in Y$, $z_1, z_2 \in Z^*$ such that $w = x_1 y_1 z_1 = x_2 y_2 z_2$ and $x_1 \neq x_2$, $z_1 \neq z_2$. From this it follows that there exist $\alpha_1 \in X^*$ and $\alpha_2 \in (X \cap Z)^*$ such that either $x_1 = x_2 \alpha_1 y_2 \alpha_2$ or $x_2 = x_1 \alpha_1 y_1 \alpha_2$. Hence either $y_2 \in X \cap Y$ and $y_1 \in Y \cap Z$ or $y_1 \in X \cap Y$ and $y_2 \in Y \cap Z$. Both cases give rise to the conclusion that $X \cap Y \neq \varnothing$ and $Y \cap Z \neq \varnothing$. On the other hand if $X \cap Y \neq \varnothing$ and $Y \cap Z \neq \varnothing$, then for all $w \in X^*(X \cap Y)(Y \cap Z)Z^*$ there exist v_1, $v_2 \in K$ such that $v_1 \neq v_2$ and *iden* $v_1 =$ *iden* $v_2 = w$. This proves the theorem. ∎

COROLLARY IV.3. *A* 1*S grammar H, with sel* $H = X^*\bar{Y}Z^*$ *is selector deterministic if and only if* $X \cap Y = \varnothing$ *or* $Y \cap Z = \varnothing$.

It turns out that the restriction of global determinism imposed on 1S grammars provides a characterization of context-free languages. (We would like to remind the reader that we do not know whether or not $\mathscr{L}(1\mathrm{S}) = \mathscr{L}(\mathrm{CF})$, hence we do not know whether global determinism constitutes a normal form for 1S grammars.)

THEOREM IV.4. *A language is context-free if and only if it is generated by a globally deterministic* 1*S grammar.*

Proof. Let $H=(\Sigma, h, S, \Delta, X^*\bar{Y}Z^*)$ be a globally deterministic 1S grammar. We assume that $X \cap Y = \varnothing$. (The case $Y \cap Z = \varnothing$ can be dealt with analogously). According to Lemma III.3 we may assume that $\Sigma \setminus \Delta = Y$, $h(a) = \{a\}$ for all $a \in \Delta$ and $a \in h(a)$ for all $a \in \Sigma$. (The constructions used in the proofs of Lemmas III.2 and III.3 do not violate the condition $X \cap Y = \varnothing$). Since $X \cap Y = \varnothing$, $X \subseteq \Delta$ and we may assume that if $\alpha \in h(a)$ for some $a \in Y$, then $\alpha \in \Delta^*$ or $\alpha \in X^*Y(((X \cap Z) \setminus Y)^* (Y \cap Z))^*(Z \cap \Delta)^*$.

It follows from Lemma III.1 that if u and v are as described in the statement of this lemma, then either $v \in (\Sigma \setminus Y)^*$, which implies $v \in L(H)$, or $u = xyz_1 y_1 z_2$, with $x \in X^*$, $y \in Y$, $y_1 \in Y \cap Z$ and z_1, $z_2 \in Z^*$, and $v = x\alpha z_1 y_1 z_2$ with $\alpha \in h(y)$ and $\alpha \in \Delta^*(\Delta \setminus X)\Delta^*$.

We will construct now a 1S grammar H_1, equivalent with H such that this last case cannot occur in H_1.

Let $\dot{Y} = \{\dot{a} : a \in Y\}$ and $\hat{Y} = \{\hat{a} : a \in Y\}$ be new mutually disjoint alphabets. Let $\Gamma = \Delta \cup \dot{Y} \cup \hat{Y} \cup \{S\}$. (We assume that $S \notin \Delta$; if $S \in \Delta$, then $L(H) = \{S\} \in \mathscr{L}(\mathrm{CF})$.) Let g be the finite substitution from Γ^* into Γ^*, which is defined in the following way.

$$g(S) = \{\hat{S}\}.$$

Let $\alpha \in h(a)$, for some $a \in Y$. If $\alpha = \alpha_1 y_1 \alpha_2 y_2 \alpha_3 \cdots \alpha_{k-1} y_{k-1} \alpha_k y_k \alpha_{k+1}$, where $\alpha_1 \in X^*$, $\alpha_2,\dots,\alpha_k \in (X \cap Z)^*$, $\alpha_{k+1} \in (Z \cap \Delta)^*$, $y_i \in Y$, $1 \leqslant i \leqslant k$, then $\alpha_1 \dot{y}_1 \alpha_2 \dot{y}_2 \alpha_3 \cdots \alpha_{k-1} \dot{y}_{k-1} \alpha_k \hat{y}_k \alpha_{k+1} \in g(\hat{a})$ and $\alpha_1 \dot{y}_1 \alpha_2 \dot{y}_2 \alpha_3 \cdots \alpha_{k-1} \dot{y}_{k-1} \alpha_k \dot{y}_k \alpha_{k+1} \in g(\dot{a})$. If $\alpha \in X^*$, then $\alpha \in g(\hat{a})$ and $\alpha \in g(\dot{a})$. If $\alpha \in \Delta^*(\Delta \setminus X)\Delta^*$, then $\alpha \in g(\hat{a})$ and $\alpha \notin g(\dot{a})$. For all $a \in \Delta$, $g(a) = \{a\}$.

Let $Y_1 = \dot{Y} \cup \hat{Y}$ and $Z_1 = (Z \setminus Y) \cup \{\dot{a} : a \in Z \cap Y\} \cup \{\hat{a} : a \in Z \cap Y\}$. $H_1 = (\Gamma, g, S, \Delta, X^* Y_1 Z_1^*)$.

Obviously $L(H) = L(H_1)$ and in any derivation step in H_1 the left most non-terminal symbol is rewritten and a production with right-hand side $\alpha \in \Delta^*(\Delta \setminus X)\Delta^*$ can only be applied to the right most non-terminal symbol. Let $G = (\Gamma, g, S, \Delta)$. From the above it follows that the set of all words in $L(G)$ obtained by left most derivations equals $L(H_1)$. Hence $L(G) = L(H_1)$ (cf. Salomaa 1973). Since G is an $E0S$ system and $L(H_1) = L(H)$, we conclude that $L(H) \in \mathscr{L}(\mathrm{CF})$.

On the other hand, let L be a context-free language, then there exists an $E0S$ system $G_1 = (\Sigma, h, S, \Delta)$, such that $L(G_1) = L$ and, for all $a \in \Delta$, $h(a) = \{a\}$.

Clearly $H_2 = (\Sigma, h, S, \Delta, \Delta^* \overline{\Sigma \setminus \Delta}\, \Sigma^*)$ is a globally deterministic $1S$ grammar, which generates $L(G_1) = L(G)$. Hence the theorem holds. ∎

The characterization provided by Theorem IV.3 gives rise to the following definition.

DEFINITION IV.2. Let H be an S grammar and let *sel* $H = \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^*$ H is called a *left-sided disjoint S grammar* if, for all $i \in \{1,\dots,n\}$, $X_i \cap Y_i = \varnothing$. H is called a *right-sided disjoint S grammar* if, for all $i \in \{1,\dots,n\}$, $Y_i \cap Z_i = \varnothing$. ∎

It turns out that both, left-sided disjointness and right-sided disjointness, constitute a normal form for S grammars.

THEOREM IV.5. *For every S grammar there exists an equivalent left-sided disjoint S grammar and an equivalent right-sided disjoint S grammar.*

Proof. Let $H = (\Sigma, h, S, \Delta, \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^*)$ be an arbitrary S grammar. We will show the existence of an equivalent left-sided disjoint S grammar; the existence of an equivalent right-sided disjoint S grammar can be shown analogously.

Let, for $i \in \{1,\dots,n\}$, $X_{\langle i \rangle} = \{a_{\langle i \rangle} : a \in X_i\}$, $\dot{Y}_{\langle i \rangle} = \{\dot{a}_{\langle i \rangle} : a \in Y_i\}$ and $Z_{\langle i \rangle} = \{a_{\langle i \rangle} : a \in Z_i\}$ be new alphabets such that $\bigcup_{i=1}^n (X_{\langle i \rangle} \cup Z_{\langle i \rangle}) \cap \bigcup_{i=1}^n \dot{Y}_{\langle i \rangle} = \varnothing$ and if $i \neq j$, $1 \leqslant i, j \leqslant n$, then $X_{\langle i \rangle} \cap X_{\langle j \rangle} = \varnothing$, $\dot{Y}_{\langle i \rangle} \cap \dot{Y}_{\langle j \rangle} = \varnothing$ and $Z_{\langle i \rangle} \cap Z_{\langle j \rangle} = \varnothing$.

Let $\hat{\Sigma} = \bigcup_{i=1}^n (X_{\langle i \rangle} \cup \dot{Y}_{\langle i \rangle} \cup Z_{\langle i \rangle})$ and let $\check{\Sigma} = \{\check{a} : a \in \Sigma\}$. $\check{\Sigma} \cap (\Sigma \cup \hat{\Sigma}) = \varnothing$.

Let $\Gamma = \Sigma \cup \hat{\Sigma} \cup \check{\Sigma}$. The finite substitution g from Γ^* into Γ^* is defined as follows. If $a \in \Sigma \setminus \bigcup_{i=1}^n (X_i \cup Y_i \cup Z_i)$, then $g(a) = \{a\}$. For $i \in \{1,..., n\}$, if $a \in X_i \cup Z_i$, then $a_{\langle i \rangle} \in g(a)$ and $g(a_{\langle i \rangle}) = \{a\}$, if $a \in Y_i$ and $a_1 \cdots a_k \in h(a)$, $a_j \in \Sigma$, $1 \leqslant j \leqslant k$, then $\dot{a}_{\langle i \rangle} \in g(a)$ and $\check{a}_1 \cdots \underline{\check{a}}_k \in g(\dot{a}_{\langle i \rangle})$, if $\underline{\check{a}} \in \check{\Sigma}$, then $g(\check{a}) = \{a\}$. Let $K = \hat{\Sigma}^* \bar{\Sigma} \Sigma^* \cup \bigcup_{i=1}^n X_{\langle i \rangle}^* \dot{\bar{Y}}_{\langle i \rangle} Z_{\langle i \rangle}^* \cup \Sigma^* (\bar{\hat{\Sigma}} \cup \bar{\check{\Sigma}})(\hat{\Sigma} \cup \check{\Sigma})^*$. Obviously, the S grammar $H' = (\Gamma, g, S, \Delta, K')$ is left-sided disjoint. It is easy to see that $L(H') = L(H)$. Hence the theorem holds. ▮

As an immediate consequence of the construction from the proof of Theorem IV.5 we get even a stronger normal form for S grammars.

THEOREM IV.6. *For every S grammar, there exists an equivalent selector deterministic left-sided disjoint S grammar and an equivalent selectors deterministic right-sided disjoint S grammar.*

We have seen that left- or right-continuous grammars constitute a normal form for continuous grammars. The situation is quite different in the case of sequential grammars.

DEFINITION IV.3. Let H be an S grammar and let *sel* $H = \bigcup_{i=1}^n X_i^* \bar{Y}_i Z_i^*$. H is called a *left-sequential* grammar (abbreviated *$l - S$ grammar*) if, for all $i \in \{1,..., n\}, X_i = \varnothing$. H is called a *right-sequential* grammar (abbreviated *$r - S$ grammar*) if, for all $i \in \{1,..., n\}$, $Z_i = \varnothing$. ▮

We will denote the classes of languages generated by $l - S$ and $r - S$ grammars as $\mathscr{L}(l - \mathrm{S})$ and $\mathscr{L}(r - \mathrm{S})$, respectively.

It turns out that considering only left- (or right-) sequential grammars restricts considerably the class of generated languages.

THEOREM IV.7. $\mathscr{L}(l - \mathrm{S}) = \mathscr{L}(r - \mathrm{S}) = \mathscr{L}(\mathrm{Reg})$.

Proof. Let $G = (\Sigma, P, S, \Delta)$ be a right-linear (left-linear) grammar. Let h_p be the finite substitution from Γ^* into Γ^* defined by $h_p(a) = \{a\}$, for all $a \in \Delta$, and $\alpha \in h_p(a)$ if and only if $a \to \alpha \in P$, for all $a \in \Sigma \setminus \Delta$. Obviously the right- sequential (left-sequential) grammar $(\Sigma, h_p, S, \Delta, \Sigma^* \bar{\Sigma})$ $((\Sigma, h_p, S, \Delta, \bar{\Sigma} \Sigma^*))$ generates $L(G)$. Hence $\mathscr{L}(\mathrm{Reg}) \subseteq \mathscr{L}(r - S)$ and $\mathscr{L}(\mathrm{Reg}) \subseteq \mathscr{L}(l - S)$.

We will prove that $\mathscr{L}(r - S) \subseteq \mathscr{L}(\mathrm{Reg})$ as follows. (That $\mathscr{L}(l - S) \subseteq \mathscr{L}(\mathrm{Reg})$ can be shown analogously.)

Let $H = (\Sigma, h, S, \Delta, \bigcup_{i=1}^n X_i^* \bar{Y}_i)$ be an arbitrary right-sequential grammar. Without loss of generality we may assume that for all $a \in \Delta$, $h(a) = \{a\}$, for all $a \in \Sigma \setminus \Delta$, $h(a) \subseteq \Delta^* (\Sigma \setminus \Delta)^*$ and there exists an $i \in \{1,..., n\}$ such that $S \in Y_i$.

The $E0S$ system $G = (\Gamma, g, S, \Delta)$ is defined in the following way.

$\Gamma = \Delta \cup \{S\} \cup \Xi$, where $\Xi = \{[a, V] : a \in \Sigma \setminus \Delta,\ V \subseteq \Sigma\}$ and Ξ, $\{S\}$ and

Δ are mutually disjoint. (If $S \in \Delta$, then $L(H) = \{S\}$ and then $L(H) \in \mathscr{L}(\mathrm{Reg})$ trivially holds.)

g is defined as follows.

For all $a \in \Gamma$, $a \in g(a)$.

$$g(S) \supseteq (h(S) \cap \Delta^*) \cup \{\alpha[A_1, alph\ \alpha] \cdots [A_l, alph\ \alpha A_1 \cdots A_{l-1}]: \alpha A_1 \cdots A_l \in h(S),\ \alpha \in \Delta^*,\ l \geqslant 1,\ A_j \in \Sigma \setminus \Delta,\ 1 \leqslant j \leqslant l\}.$$

For all $[a, V] \in \Xi$, such that $V \cap (\Sigma \setminus \Delta) \neq \varnothing$ and such that $a \in Y_i$ and $V \subseteq X_i$ for some $i \in \{1,..., n\}$,

$$g([a, V]) \supseteq \{[A_1, V][A_2, V \cup \{A_1\}] \cdots [A_l, V \cup \{A_1,..., A_{l-1}\}]: A_1 ... A_l \in h(a),\ l \geqslant 0,\ A_j \in \Sigma \setminus \Delta,\ 1 \leqslant j \leqslant l\}.$$

For all $[a, V] \in \Xi$, such that $V \cap (\Sigma \setminus \Delta) = \varnothing$ and such that $a \in Y_i$ and $V \subseteq X_i$ for some $i \in \{1,..., n\}$, $g([a, V]) \supseteq \{\alpha[A_1, V \cup alph\ \alpha] \cdots [A_l, V \cup alph\ \alpha A_1 \cdots A_{l-1}]: \alpha A_1 \cdots A_l \in h(a),\ \alpha \in \Delta^*,\ l \geqslant 0,\ A_j \in \Sigma \setminus \Delta,\ 1 \leqslant j \leqslant l\}$. It is easy to see that $L(G) = L(H)$.

Moreover, a right-linear grammar G', equivalent with G, can be constructed from G, by iterating the following procedure.

For all $[a, V] \in \Xi$, such that $g([a, V]) = \{[a, V]\}$, remove all productions from g in which $[a, V]$ occurs at the right-hand side. If this has been done, replace all productions with right-hand side $\alpha[A_1, V] \cdots [A_l, V \cup \{A_1,..., A_{l-1}\}]$ by $\alpha[A_1, V]$ and remove $g(a) = \{a\}$ for all $a \in \Delta$. It is not difficult to see that $L(G') = L(G) = L(H)$. Clearly G' is a right-linear grammar. (If G' has erasing productions they can be removed in a standard way, see, e.g., Salomaa, 1973.) Hence $\mathscr{L}(r - \mathrm{S}) \subseteq \mathscr{L}(\mathrm{Reg})$ and the theorem holds. ∎

One can relax the left- (or right-) sequential restriction on S grammars by allowing an S grammar to have selectors of the type $\bar{Y}Z^*$ as well as selectors of the type $X^*\bar{Y}$.

DEFINITION IV.4. Let H be an S grammar and let $sel\ H = \bigcup_{i=1}^{n} X_i^* \bar{Y}_i Z_i^*$. H is called a *one-sided sequential grammar* (abbreviated *$s - S$ grammar*) if, for all $i \in \{1,..., n\}$, $X_i = \varnothing$ or $Z_i = \varnothing$. ∎

We will denote the class of languages generated by $s - S$ grammars as $\mathscr{L}(s - S)$.

We will demonstrate now that propagating one-sided sequential grammars generate precisely the class of linear languages.

First, we need some terminology.

DEFINITION IV.5. Let $H = (\Sigma, h, S, \Delta, \bigcup_{i=1}^{n} X_i^* \bar{Y}_i Z_i^*)$ be an S grammar, let $K_i = X_i^* \bar{Y}_i Z_i^*$, $1 \leqslant i \leqslant n$.

If for some $i \in \{1,\ldots,n\}$, $X_i = \emptyset (Z_i = \emptyset)$, then K_i is called a *left-selector* (respectively *right-selector*).

If K_i, for some $i \in \{1,\ldots,n\}$ is a left-selector as well as a right-selector, then K_i is called *narrow*.

Let K_i, for some $i \in \{1,\ldots,n\}$, be a left-selector (respectively right-selector).

If $y \in Y_i$, then y is called a *left-symbol* (resp. *right-symbol*).

If $Y_i = \{y\}$, $y \in \Sigma$, then K_i is called a *y-left selector* (*y-right selector*) and we write $y = leftm(K_i)$ (resp. $y = rightm(K_i)$). ∎

LEMMA IV.1. *For every propagating $s - S$ grammar there exists an equivalent linear grammar.*

Proof. Let $H = (\Sigma, h, S, \Delta, K)$ be a propagating S grammar, with $K = \bigcup_{i=1}^{n} K_{l,i} \cup \bigcup_{i=1}^{m} K_{r,i} \cup \{\bar{S}\}$.

Each $K_{l,i}$ is a left-selector and each $K_{r,i}$ is a right-selector. Obviously we may assume that S does not appear on the right-hand side of any production in h, nor does S appear as a left-symbol or as a right-symbol in $\bigcup_{i=1}^{n} K_{l,i} \cup \bigcup_{i=1}^{m} K_{r,i}$; all $K_{l,i}$ and $K_{r,j}$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$, contain only one left-symbol or right-symbol respectively and none of them is narrow.

First we will construct an $s - S$ grammar G such that $L(G) = ¢\, L(H)\, \$$, where $¢, \$ \notin \Sigma$, in G only non-terminals can be rewritten and all productions except the productions for the axiom of G are either left-linear or right-linear. Let, for each $a \in \Delta$, $\Xi_a = \{M_{l,a}, N_{l,a}, M_{r,a}, N_{r,a}\}$ be a set of new symbols, such that $\Xi_a \cap \Xi_b = \emptyset$, if $a \neq b$, and let F, $¢$ and $\$$ be new symbols. $\Gamma = \Sigma \cup \bigcup_{a \in \Delta} \Xi_a \cup \{F, ¢, \$\}$.

Let g be the finite substitution from Γ^* into Γ^* which is defined as follows.

(1) For all $a \in \Gamma$, $F \in g(a)$.

(2) For all $a \in \Delta$, $¢a \in g(M_{l,a})$ and $a\$ \in g(M_{r,a})$.

(3) For all $y \in \Sigma \setminus \Delta$, such that K contains a y-left selector and $a\alpha \in h(y)$ for some $\alpha \in \Delta^*$.

(i) $a\alpha \in g(y)$ if $a \in \Sigma \setminus \Delta$ and

(ii) $\{M_{l,a}\alpha, N_{l,a}\alpha\} \subseteq g(y)$ if $a \in \Delta$.

(4) For all $y \in \Sigma \setminus \Delta$, such that K contains a y-right selector and $\alpha a \in h(y)$ for some $\alpha \in \Delta^*$.

(i) $\alpha a \in g(y)$ if $a \in \Sigma \setminus \Delta$ and

(ii) $\{\alpha M_{r,a}, \alpha N_{r,a}\} \subseteq g(y)$ if $a \in \Delta$.

(5) For all $y \in \Delta$, such that K contains a y-left selector and $a\alpha \in h(y)$, for some $\alpha \in \Delta^*$

(i) $a\alpha \in g(N_{l,y})$ if $a \in \Sigma \setminus \Delta$ and

(ii) $\{M_{l,a}\alpha, N_{l,a}\alpha\} \subseteq g(N_{l,y})$ if $a \in \Delta$.

(6) For all $y \in \Delta$, such that K contains a y-right selector and $\alpha a \in h(y)$, for some $\alpha \in \Delta^*$

(i) $\alpha a \in g(N_{r,y})$ if $a \in \Sigma \setminus \Delta$ and

(ii) $\{\alpha M_{r,a}, \alpha N_{r,a}\} \subseteq g(N_{r,y})$ if $a \in \Delta$.

(7) Let U be the set of all sentential forms w of H such that $|w| \geqslant 2$ and w can be derived in H in such a way that all intermediate words are of length 1.

Let $one(H) = \{w \in L(H) : |w| = 1\}$.

Let φ_l and φ_r be the finite substitutions on Σ^* defined by $\varphi_l(a) = \varphi_r(a) = \{a\}$ if $a \in \Sigma \setminus \Delta$ and $\varphi_l(a) = \{M_{l,a}, N_{l,a}\}$, $\varphi_r(a) = \{M_{r,a}, N_{r,a}\}$ if $a \in \Delta$. Then let $W = \{\varphi_l(a)\alpha\, \varphi_r(b) : \alpha \in \Delta^*,\ a\alpha b \in U\}$.

$$g(S) \supseteq W \cup \{¢w\$: w \in one(H)\}.$$

Let T be the smallest set of selectors containing the following languages. $\{\bar{S}\} \subseteq T$.

If, for some $i \in \{1,..., m\}$, $K_{r,i} = X^*\{\bar{y}\}$, then

(i) $(X \cup \varphi_l(X))^*\{\bar{y}\} \subseteq T$, if $y \in \Sigma \setminus \Delta$ and

(ii) $(X \cup \varphi_l(X))^*\{\bar{N}_{r,y}\} \subseteq T$, if $y \in \Delta$.

If, for some $i \in \{1,..., n\}$, $K_{l,i} = \{\bar{y}\}\, Z^*$, then

(i) $\{\bar{y}\}(Z \cup \varphi_r(Z))^* \subseteq T$, if $y \in \Sigma \setminus \Delta$ and

(ii) $\{\bar{N}_{l,y}\}(Z \cup \varphi_r(Z))^* \subseteq T$ if $y \in \Delta$.

For every $y \in \Delta$, $\Gamma^*\{\bar{M}_{r,y}\} \subseteq T$ and $\{\bar{M}_{l,y}\}\, \Gamma^* \subseteq T$. Let $G = (\Gamma, g, S, \Delta \cup \{¢, \$\}, T)$. Clearly $L(G) = ¢\, L(H)\$$ and all productions in G, which are not S-productions are either left- or right-linear. Clearly $T = \{\bar{S}\} \cup \bigcup_{i=1}^{u} T_{l,i} \cup \bigcup_{i=1}^{s} T_{r,i}$ for some u and s, u, $s \geqslant 1$, where $T_{l,i}$ is a left-selector, $1 \leqslant i \leqslant u$, and $T_{r,i}$ is a right-selector, $1 \leqslant i \leqslant s$. Moreover (by proper indexing of the productions) we can assume that $\bigcup_{i=1}^{u} leftm(T_{l,i}) \cap \bigcup_{i=1}^{s} rightm(T_{r,i}) = \varnothing$ and if $i \neq j$, then $leftm(T_{l,i}) \neq leftm(T_{l,j})$, $1 \leqslant i,j \leqslant u$, and $rightm(T_{r,i}) \neq rightm(T_{r,j})$, $1 \leqslant i$, $j \leqslant s$. Thus for every left-symbol or right-symbol y, there is a unique selector, denoted $sel\ y$, which "uses" this production.

Let, for all $i \in \{1,..., u\}$, $T_{l,i} = \bar{Y}_{l,i} Z_{l,i}^*$ and, for all $i \in \{1,..., s\}$, $T_{r,i} = X_{r,i}^* \bar{Y}_{r,i}$.

Let $T_l = \bigcup_{i=1}^{u} T_{l,i}$, $T_r = \bigcup_{i=1}^{s} T_{r,i}$, $Y_l = \bigcup_{i=1}^{u} Y_{l,i}$, $Y_r = \bigcup_{i=1}^{s} Y_{r,i}$ and $Y = Y_l \cup Y_r$. For every $i \in \{1,..., u\}$, let $forb(T_{l,i}) = \{y : \{y\} = Y_{r,j}$, $1 \leqslant j \leqslant s$ and $y \notin Z_{l,i}\}$ and for every $i \in \{1,..., s\}$, let $forb(T_{r,i}) = \{y : \{y\} = Y_{l,j}$,

$1 \leqslant j \leqslant u$ and $y \notin X_{r,i}\}$. If π is a production of the form $y \to \alpha$, where $y \in Y$, then $forb(\pi) = \{T_{l,i}: (\exists a)_\Delta[a \in alph \quad \alpha \quad \text{and} \quad a \notin Z_{l,i}], \quad 1 \leqslant i \leqslant u\} \cup \{T_{r,i}: (\exists a)_\Delta[a \in alph\ \alpha \text{ and } a \notin X_{r,i}], i \leqslant i \leqslant s\}$.

Let G_1 be the linear grammar, constructed as follows.

$$G_1 = (V, P, Z, V_T),\ V = V_N \cup V_T,\ V_N \cap V_T = \varnothing.$$

$$V_N = \{[v_1, v_2, v_3, v_4] : v_1 \in \Delta^* \text{ and } |v_1| \leqslant maxr(G) - 2,$$
$$v_2 \in Y_l \cup \{¢\},\ v_3 \in Y_r \cup \{\$\},\ v_4 \subseteq T \setminus \{\bar{S}\}\} \cup \{Z\}.$$
$$V_T = \Delta.$$

P is defined in the following way.

I. The Z-productions in P are

$Z \to x$, for all $x \in one(H)$;

$Z \to [\alpha, ¢, \$, \varnothing]$, for every α such that $c\alpha d \in W$ for some $c, d \in \Gamma$.

II. The productions for the other non-terminals of G_1 are $[v_1, v_2, v_3, v_4] \to \alpha[v_1, v_2', v_3, v_4']$, if there exists a left-linear production

$\pi : v_2' \to v_2\alpha$ in G such that $v_3 \notin forb\ (sel\ v_2'), forb\ (\pi) \cap v_4 = \varnothing$ and

$v_4' = v_4 \cup sel\ v_2'$;

$[v_1, v_2, v_3, v_4] \to [v_1, v_2, v_3', v_4]\alpha$, if there exists a right-linear production

$\pi : v_3' \to \alpha v_3$ in G such that $v_2 \notin forb\ (sel\ v_3'), forb\ (\pi) \cap v_4 = \varnothing$

and $v_4' = v_4 \cup sel\ v_3'$;

$[v_1, v_2, v_3, v_4] \to v_1$ if $v_2 v_1 v_3 \in W$.

By observing that in G_1 bottom-up simulations of top-down derivations in G are performed, it is not difficult to see that $L(G_1) = L(G)$. ∎

LEMMA IV.2. *For every linear grammar there exists an equivalent $s - S$ grammar.*

Proof. Let $G = (\Sigma, P, S, \Delta)$ be a linear grammar. Without loss of generality we assume the following. $L(G) = L(G_1) \cup L(G_2) \cup L(G_3)$, where $G_1 = (\Sigma_1, P_1, S_1, \Delta)$ is a right-linear grammar, $G_2 = (\Sigma_2, P_2, S_2, \Delta)$ is a left-linear grammar, $G_3 = (\Sigma_3, P_3, S_3, \Delta)$ is a linear grammar, such that $S_3 \to \alpha$ in P_3 implies that $\alpha \in \Delta^+(\Sigma_3 \setminus (\Delta \cup \{S\}))\Delta^+$ and S_3 never occurs at the right-hand side of any production of G_3. $(\Sigma_1 \setminus \Delta) \cap (\Sigma_2 \setminus \Delta) \cap (\Sigma_3 \setminus \Delta) = \varnothing$ and $S \to S_1, S \to S_2, S \to S_3$ are the only S-productions of G.

We will first construct an $s - S$ grammar for G_3.

Let $P_3 = P'_3 \cup P''_3$, where $P'_3 = \{\pi : \pi = A \to \alpha B \beta,\ A,\ B \in \Sigma \setminus \Delta,\ \alpha,\ \beta \in \Delta^*\}$ and $P''_3 = \{\pi : \pi = A \to \alpha, \alpha \in \Delta^*\}$. If $\pi = A \to \alpha B \beta \in P'_3$ we denote $lh\pi = A$, $rh\pi = B$, $lrh\pi = \alpha$ and $rrh\pi = \beta$; if $\pi = A \to \alpha \in P''_3$ we denote $lh\pi = A$ and $rh\pi = \alpha$. Let $\Pi_l = \{[l, \pi] : \pi \in P'_3\}$, $\Pi'_l = \{[l, \pi]' : \pi \in P'_3\}$, $\Pi_r = \{[\pi, r] : \pi \in P'_3\}$ and $\Pi'_r = \{[\pi, r]' : \pi \in P'_3\}$ be new, mutually disjoint alphabets.

Let Z be a new symbol.

$$\Gamma = \Pi_l \cup \Pi'_l \cup \Pi_r \cup \Pi'_r \cup \Delta \cup \{Z\}.$$

The finite substitution g from Γ^* into Γ^* is defined as follows. $g(Z) = \{[l, \pi]'\, \alpha [\pi, r]' : \pi \in P'_3,\ \tau \in P''_3$, such that $rh\tau = \alpha$ and $lh\tau = rh\pi\}$. For all $[l, \pi] \in \Pi_l$, such that $lh\pi \neq S_3$, $g([l, \pi]) = \{[l, \tau]'\alpha : \alpha = lrh\pi$ and $\tau \in P'_3$ such that $lh\pi = rh\tau\}$. For all $[\pi, r] \in \Pi_r$, such that $lh\pi \neq S_3$, $g([\pi, r]) = \{\alpha[\tau, r]' : \alpha = rrh\pi$ and $\tau \in P'_3$ such that $lh\pi = rh\tau\}$. For all $\pi \in P'_3$ such that $lh\pi = S_3$, $g([l, \pi]) = \{\alpha : \alpha = lrh\pi\}$ and $g([\pi, r]) = \{\alpha : \alpha = rrh\,\pi\}$. For all $[l, \pi]' \in \Pi'_l$, $g([l, \pi]') = \{[l, \pi]\}$ and for all $[\pi, r]' \in \Pi'_r\, g([\pi, r]') = \{[\pi, r]\}$. For all $a \in \Delta$, $g(a) = \{a\}$.

$H_3 = (\Gamma, g, Z, \Delta, K_3)$, with

$$K_3 = \{\bar{Z}\} \cup \bigcup_{\pi \in P'_3} (\{\overline{[l, \pi]'}\}(\Delta \cup \{[\pi, r]')^* \cup (\Delta \cup [l, \pi])^* \{[\overline{\pi, r}]'\}$$
$$\cup \{[\overline{l, \pi}]\}(\Delta \cup \{[\pi, r]\})^*) \cup (\Delta \cup \Pi'_l)^* \left(\bigcup_{\pi \in P'_3} [\overline{\pi, r}] \right)$$

It is rather straight forward to prove that $L(H_3) = L(G_3)$. (H_3 simulates G_3 in a bottom-up way.)

Let $H_1 = (\Sigma_1, h_{p_1}, S_1, \Delta, \Sigma_1^* \bar{\Sigma}_1)$ and $H_2 = (\Sigma_2, h_{p_2}, S_2, \Delta, \bar{\Sigma}_2 \Sigma_2^*)$ be the $r - S$ and $l - S$ grammar, constructed as in the proof of Theorem IV.7, which are equivalent with G_1 and G_2.

Define $H = (\Xi, h, \hat{Z}, \Delta, K)$ by $\Xi = \Sigma_1 \cup \Sigma_2 \cup \Gamma \cup \{\hat{Z}\}$, $h(\hat{Z}) = \{S_1, S_2, Z\}$, $h(a) = \{a\}$ for all $a \in \Delta$, $h(a) = h_{p_1}(a)$ for all $a \in \Sigma_1 \setminus \Delta$, $h(a) = h_{p_2}(a)$ for all $a \in \Sigma_2 \setminus \Delta$, and $h(a) = g(a)$ for all $a \in \Gamma \setminus \Delta$.

Let $K = \{\bar{\hat{Z}}\} \cup \Sigma_1^* \bar{\Sigma}_1 \cup \bar{\Sigma}_2 \Sigma_2^* \cup K_3$.

Then H is an $s - S$ grammar and obviously $L(H) = L(G)$.

Hence the lemma holds. ∎

THEOREM IV.8. *A language is linear if and only if it is generated by a propagating one-sided sequential grammar.*

Proof. Immediate from Lemmas IV.1 and IV.2 when it is observed that in the proof of Lemma IV.2 no erasing productions are introduced. ∎

Note that the main construction we use for the proof of Lemma IV.1 is different from the construction used in Theorem IV.7 (the former is a bottom-up simulation, while the latter is a top-down simulation). Moreover the construction from the proof of Theorem IV.7 can handle the case of erasing, while the construction from Lemma IV.1 *cannot* handle erasing productions.

As a matter of fact allowing erasing in one-sided sequential grammars increases their language generating power.

THEOREM IV.9. $\mathscr{L}(s-\mathrm{S}) \supsetneq \mathscr{L}(\mathrm{Lin})$.

Proof. The inclusion is a consequence of Theorem IV.8. Its properness is demonstrated by the construction of an $s-S$ grammar, which generates the non-linear language $\{¢\, d^k a^m * b^n : m > n \geqslant 0, \text{ and } k > m-n\}$. That this is a non-linear language follows from the well-known pumping lemma for linear languages (see, e.g., Berstel (1979)). However, notice that this language is context-free.

Let $H = (\Sigma, h, S, \Delta, K)$ be defined as follows.

$\Sigma = \{S, L, L_1, L_2, L_3, L_4, R, R_1, B, C, M, M_1, M_2, a, b, d, *, ¢\}$.

$\Delta = \{a, b, d, *, ¢\}$.

$h(S) = \{L * R\}$, $h(L) = \{L_1 a\}$, $h(L_1) = \{L, L_2\}$, $h(L_2) = \{L_3 d\}$,

$h(L_3) = \{L_4\}$, $h(L_4) = \{L_3 d, ¢\}$, $h(R) = \{bR_1, BCM_1\}$, $h(R_1) = \{R\}$,

$h(M_1) = \{M, M_2\}$, $h(M) = \{BCM_1\}$, $h(M_2) = \{\Lambda\}$, $h(B) = \{\Lambda\}$, $h(C) = \{\Lambda\}$,

$h(a) = \{a\}$, $h(b) = \{b\}$, $h(d) = \{d\}$, $h(*) = \{*\}$ and $h(¢) = \{¢\}$.

$$\begin{aligned}K = {} & \{\bar{S}\} \cup \{\bar{L}\}\{a, b, *, R_1, B, C, M_1\}^* \cup \{\bar{L}_1\}\{a, b, *, R, B, C, M, M_2\}^* \\ & \cup \{\bar{L}_2\}\{a, b, *, B, C\}^* \cup \{\bar{L}_3\}\{a, b, *, d, B, C\}^* \cup \{\bar{L}_4\}\{a, b, *, d, B, C\}^* \\ & \cup \{a, b, *, L\}^*\{\bar{R}\} \cup \{a, b, *, L_1\}^*\{\bar{R}_1\} \cup \{a, b, *, L_1, B, C\}^*\{\bar{M}_1\} \\ & \cup \{a, b, *, L_2, B, C\}^*\{\bar{M}_2\} \cup \{a, b, *, L, B, C\}^*\{\bar{M}\} \\ & \cup \{a, b, *, d, L_4\}^*\{\bar{B}\} \cup \{a, b, *, d, L_3, B, C\}^* \{\bar{C}\}.\end{aligned}$$

A typical derivation in H is of the following form. (The underlining of the symbols denotes which symbol will be rewritten in the next derivation step.)

$$\underline{S} \Rightarrow L * \underline{R} \Rightarrow \underline{L} * bR_1 \Rightarrow L_1 a * b\underline{R}_1 \Rightarrow \underline{L}_1 a * bR \Rightarrow La * b\underline{R}.$$

Obviously this "cycle" can be repeated, yielding, $n \geqslant 0$, $La^n * b^n\underline{R}$. Since rewriting L_1 into L_2 yields a word for which K does not provide a selector word, the only way to leave this cycle is to apply another production to R.

Hence $La^n * b^n\underline{R} \Rightarrow \underline{L}a^n * b^nBCM_1 \Rightarrow L_1 a^{n+1} * b^nBC\underline{M}_1 \Rightarrow \underline{L}_1 a^{n+1} * b^nBCM \Rightarrow La^{n+1} * b^nBC\underline{M} \Rightarrow \underline{L}a^{n+1} * b^n(BC)^2M_1$.

We obtain by repeating this last four steps $m-1$ times $\underline{L}a^m a^n * b^n(BC)^{m+1}M_1$. Rewriting L_1 into L_2 yields again a blocking situation and hence if the derivation is to be successful we must apply another production to M_1. For $n \geqslant 0$, $m \geqslant 0$ this yields

$$
\begin{aligned}
&\underline{L}a^m a^n * b^n(BC)^{m+1}M_1\\
&\quad \Rightarrow L_1 a^{m+1}a^n * b^n(BC)^{m+1}\underline{M}_1 \Rightarrow \underline{L}_1 a^{m+1}a^n * b^n(BC)^{m+1}M_2\\
&\quad \Rightarrow L_2 a^{m+1}a^n * b^n(BC)^{m+1}\underline{M}_2 \Rightarrow \underline{L}_2 a^{m+1}a^n * b^n(BC)^{m+1}\\
&\quad \Rightarrow L_3 da^{m+1}a^n * b^n(BC)^{m+1}.
\end{aligned}
$$

(In this case the rewriting of L_1 into L would "block" the derivation.)

Either

$$
\begin{aligned}
\underline{L}_3 da^{m+1}a^n * b^n(BC)^{m+1} &\Rightarrow \underline{L}_4 da^{m+1}a^n * b^n(BC)^{m+1}\\
&\Rightarrow L_3 d^2a^{m+1}a^n * b^n(BC)^{m+1},
\end{aligned}
$$

or

$$
\begin{aligned}
&L_3 da^{m+1}a^n * b^n(BC)^m B\underline{C}\\
&\quad \Rightarrow \underline{L}_3 da^{m+1}a^n * b^n(BC)^m B \Rightarrow L_4 da^{m+1}a^n * b^n(BC)^m \underline{B}\\
&\quad \Rightarrow \underline{L}_4 da^{m+1}a^n * b^n(BC)^m \Rightarrow L_3 d^2a^{m+1}a^n * b^n(BC)^m,
\end{aligned}
$$

or

$$
\begin{aligned}
&L_3 da^{m+1}a^n * b^n(BC)^m B\underline{C}\\
&\quad \Rightarrow \underline{L}_3 da^{m+1}a^n * b^n(BC)^m B \Rightarrow \underline{L}_4 da^{m+1}a^n * b^n(BC)^m B\\
&\quad \Rightarrow \underline{L}_3 d^2a^{m+1}a^n * b^n(BC)^m B \Rightarrow L_4 d^2a^{m+1}a^n * b^n(BC)^m B.
\end{aligned}
$$

Hence

$$L_3 da^{m+1}a^n * b^n(BC)^{m+1} \Rightarrow^* L_4 d^l da^{m+1}a^n * b^n(BC)^m$$

where $l \geqslant 0$ and

$$\underline{L}_4 d^l da^{m+1}a^n * b^n(BC)^p \Rightarrow L_3 dd^l da^{m+1}a^n * b^n(BC)^p$$

if $p > 0$ and

$$L_4 d^l da^{m+1}a^n * b^n \Rightarrow ¢d^l da^{m+1}a^n * b^n.$$

Hence

$$L_3 da^{m+1}a^n * b^n(BC)^{m+1} \Rightarrow^* L_4 d^l da^{m+1}a^n * b^n \Rightarrow ¢d^l da^{m+1}a^n * b^n,$$

with $l \geqslant m+1$.

It is rather easy to control that all derivations in H are of this type. Hence $L(H) = \{¢d^k a^m * b^n : m > n \geqslant 0 \text{ and } k > m - n\}$. ∎

We conclude this section by demonstrating that for an arbitrary selective substitution grammar there exists an equivalent selective substitution grammar which is globally deterministic.

THEOREM IV.10. *For every E0S based s-grammar, there exists an equivalent globally deterministic E0S based s-grammar.*

Proof. Let $H = (\Sigma, h, S, \Delta, K)$ be an arbitrary $E0S$ based s-grammar.

Let $\Sigma_0 = \{a^0 : a \in \Sigma\}$, $\Sigma_1 = \{a^1 : a \in \Sigma\}$ and $\Xi = \{¢, £, \$, \rho, \sigma, Z\}$ be pairwise disjoint alphabets which are also disjoint with Σ.

Let φ be a homomorphism from $(\Sigma_0 \cup \Sigma_1)^*$ into $(\Sigma \cup \bar{\Sigma})^*$ defined by $\varphi(a^0) = a$ and $\varphi(a^1) = \bar{a}$ for all $a \in \Sigma$.

The mapping ψ from $(\Sigma_0 \cup \Sigma_1)^+$ into $\mathbb{N}$ is defined by $\psi(a_0^{i_0} \cdots a_n^{i_n}) = \sum_{j=0}^{n} i_j 2^j$, where $n \geqslant 0$ and $i_j \in \{0, 1\}$ for $0 \leqslant j \leqslant n$. (A word in $(\Sigma_0 \cup \Sigma_1)^+$ can, by means of ψ, be considered to be a binary number, written down from left to right.)

Let $\Gamma = \Sigma \cup \Sigma_0 \cup \Sigma_1 \cup \Xi$. The finite substitution g from Γ^* into Γ^* is defined as follows.

$$g(Z) = \{¢\alpha : \alpha \in h(S)\}, \qquad g(¢) = \{\rho¢, \$\}, \qquad g(\rho) = \{\sigma\},$$
$$g(\sigma) = \{\Delta\}, \qquad g(\$) = \{£\},$$
$$g(£) = \{¢, \Delta\}, \qquad \text{for all } a \in \Sigma,\ g(a) = \{a^0, a^1\},\ g(a^0) = \{a\} \text{ and } g(a^1) = h(a).$$

Let $K_0 = \rho^*\{\bar{Z}, \bar{¢}\}\,\Sigma^*$, $K_1 = \{\bar{\rho}^n \bar{\$} \bar{x} : n \geqslant 0, x \in \Sigma^+\}$ and $K_2 = \{\bar{\sigma}^n \bar{£} \bar{y} : n \geqslant 0, y \in (\Sigma_0 \cup \Sigma_1)^+$ such that $\varphi(y) \in K$ and $\psi(y) = n\}$. $H' = (\Gamma, g, Z, \Delta, K')$, with $K' = \bigcup_{j=0}^{2} K_j$.

Clearly H' generates $L(H)$, That H' is globally deterministic, can easily be seen, when it is observed that K_0, K_1 and K_2 do not have any words in common, K_0 and K_1 do not contain different words w_1 and $w_2 \in (\Gamma \cup \bar{\Gamma})^+$ such that *iden* $w_1 =$ *iden* w_2 and this property is guaranteed for K_2 by the uniqueness of the binary representation of natural numbers. Hence H' is a globally deterministic $E0S$ based s-grammar, generating $L(H)$, which proves the theorem. ∎

V. SYMMETRY AND BALANCE

Consider an S grammar $H = (\Sigma, h, S, \Delta, K)$, where K is given in the form $\{K_1, ..., K_n\}$ and $K_i = X_i^* \bar{Y}_i Z_i^*$. Rewriting a word w in H according to a K_i requires a decomposition of w in the form $w = w_1 a w_2$, where $w_1 \in X_i^*$,

$w_2 \in Z_i^*$ and $a \in Y$. In other words, this particular occurrence of a can be rewritten only if it has a "proper context" to the left and to the right of it. In this way S grammars provide a natural framework for the investigation of the role of context in rewriting systems based on sequential rewriting. Similarly one can study this problem area in the framework of C grammars.

If a component $K_i = X_i^* \bar{Y}_i Z_i^*$ of the selector of an S grammar is such that $X_i \neq Z_i$ then using such a component allows one to distinguish the "left" and the "right" side of the context of the letter under rewriting. Thus it is natural to consider the situation where such a distinction is not possible and see whether or not this affects the language generating power of the grammars considered.

In this way S grammars provide a natural framework to investigate the role of "symmetry" in rewriting systems based on sequential rewriting. Similarly one can study this problem area in the framework of C grammars.

Another way of demanding that only symmetric context is used in an S grammar is to require that whenever $X^*\bar{Y}Z^*$ is a component of the selector then so is $Z^*\bar{Y}X^*$.

This leads to the following definitions.

Definition V.1. (i) Let H be an S grammar. If $sel\ H = \bigcup_{i=1}^{n} X_i^* \bar{Y}_i X_i^*$, then H is a *symmetric nS grammar.*

(ii) Let H be a C grammar. If $sel\ H = \bigcup_{i=1}^{n} X_i^* \bar{Y}_i^+ X_i^*$, then H is a *symmetric nC grammar.*

(iii) If H is for some $n \geqslant 1$ a symmetric nS grammar or a symmetric nC grammar then we say that H is a *symmetric S grammar*, respectively a *symmetric C grammar.* ∎

The classes of languages generated by symmetric nS and symmetric nC grammars are denoted as $\mathscr{L}(\text{sym}\ nS)$ and $\mathscr{L}(\text{sym}\ nC)$, respectively. $\mathscr{L}(\text{sym}\ S) = \bigcup_{n=1}^{\infty} \mathscr{L}(\text{sym}\ nS)$ and $\mathscr{L}(\text{sym}\ C) = \bigcup_{n=1}^{\infty} \mathscr{L}(\text{sym}\ nC)$.

Definition V.2. (i) Let H be an S grammar. If $X^*\bar{Y}Z^* \subseteq sel\ H$ implies that $Z^*\bar{Y}X^* \subseteq sel\ H$, then H is a *balanced S grammar.*

(ii) Let H be a C grammar. If $X^*\bar{Y}^+Z^* \subseteq sel\ H$ implies that $Z^*\bar{Y}^+X^* \subseteq sel\ H$ then H is a *balanced C grammar.* ∎

The classes of languages generated by balanced S grammars and balanced C grammars are denoted $\mathscr{L}(\text{bal}\ S)$ and $\mathscr{L}(\text{bal}\ C)$, respectively.

It turns out that "balancing" does not influence the language generating power of neither sequential nor continuous grammars.

Lemma V.1. *For every S grammar there exists an equivalent balanced S grammar.*

Proof. Let $H=(\Sigma, h, S, \Delta, K)$ be an S grammar, with $K=\bigcup_{i=1}^{n} X_i^* \bar{Y}_i Z_i^*$. We will construct an equivalent balanced S grammar $H'=(\Gamma, g, Z, \Delta, K')$ in the following way.

$\Gamma=\Delta \cup \Delta' \cup \dot{\Sigma} \cup \{\$, Z\}$, where $\Delta'=\{a' : a \in \Delta\}$ and
$\dot{\Sigma}=\{\dot{a} : a \in \Sigma\}$ are new alphabets and \$ and Z are new symbols.
Δ, Δ', $\dot{\Sigma}$ and $\{\$, Z\}$ are pairwise disjoint.

The codings φ and ψ, respectively from Σ^* into $\dot{\Sigma}^*$ and from Σ^* into $((\dot{\Sigma} \setminus \dot{\Delta}) \cup \Delta')^*$, where $\dot{\Delta}=\{\dot{a} : a \in \Delta\}$, are defined by $\varphi(a)=\dot{a}$, for all $a \in \Sigma$, and $\psi(a)=\dot{a}$, if $a \in \Sigma \setminus \Delta$, and $\psi(a)=a'$, if $a \in \Delta$. The finite substitution g from Γ^* into Γ^* is defined as follows.

$$g(Z)=\{\$\varphi(S), \$\psi(S)\},\ g(\dot{a})=\varphi h(a) \cup \psi h(a) \text{ for all } a \in \Sigma,\ g(a')=\{a\},$$

and

$$g(a)=\{a\}, \text{ for all } a \in \Delta, \text{ and } g(\$)=\Delta.$$

$$\begin{aligned} K=\{\bar{Z}\} &\cup \bigcup_{i=1}^{n} (\varphi(X_i) \cup \psi(X_i) \cup \{\$\})^* \bar{\dot{Y}}_i (\varphi(Z_i) \cup \psi(Z_i))^* \\ &\cup \bigcup_{i=1}^{n} (\varphi(Z_i) \cup \psi(Z_i))^* \bar{\dot{Y}}_i (\varphi(X_i) \cup \psi(X_i) \cup \{\$\})^* \\ &\cup (\Delta \cup \Delta')^* (\bar{\Delta}' \cup \{\bar{\$}\}) (\Delta \cup \Delta')^*. \end{aligned}$$

To see that $L(H')=L(H)$ we make the following observations.

Since the symbol \$ is the leftmost symbol of any sentential form in which it occurs, obviously, rewriting of sentential forms using those components of K' in which \$ appears at the right-hand side of the barred symbol does not occur. Moreover if \$ is not occurring in a sentential form w of H' under rewriting, then either $w=Z$ or $w \in (\Delta \cup \Delta')^* \Delta' (\Delta \cup \Delta')^*$ in which case w can only terminate in the corresponding terminal word. Hence we conclude that any successful derivation of a word in $\$\Delta'^*$ according to H' has an analogon in the set of derivations according to H, the converse is clearly true. Since every $\$w \in \$\Delta'^*$ can only derive $v \in \Delta^*$, where $\psi(v)=w$, it follows that H and H' are equivalent.

Since H' is balanced our lemma holds. ▮

THEOREM V.1. $\mathscr{L}(\text{bal S})=\mathscr{L}(\text{S})$.

Proof. The inclusion $\mathscr{L}(\text{bal S}) \subseteq \mathscr{L}(\text{S})$ follows from Definition V.2. The converse inclusion from Lemma V.1. ▮

LEMMA V.2. *For every C grammar there exists an equivalent balanced C grammar.*

Proof. This is a straightforward adaptation of the proof of Lemma V.1. ∎

THEOREM V.2. $\mathscr{L}(\text{bal C}) = \mathscr{L}(\text{C})$.

It turns out that symmetric $1S$ grammars provide a characterization of the family of context-free languages.

THEOREM V.3. *A language is context-free if and only if it is generated by a symmetric* $1S$ *grammar.*

Proof. Let H be a symmetric $1S$ grammar; let $sel\ H = X^*\bar{Y}X^*$. Hence, by Theorem III.4, $L(H) \in \mathscr{L}(\text{CF})$.

Conversely, every context-free language generated by an $E0S$ system $G = (\Sigma, h, S, \Delta)$ is also generated by the $1S$ grammar $H = (\Sigma, h, S, \Delta, \Sigma^*\bar{\Sigma}\Sigma^*)$. ∎

The symmetric restriction imposed on S grammars provides a characterization of the class of forbidding grammars (see, for example, Penttonen (1975) and Lomkovskaja (1972)).

First we recall the definition of a forbidding grammar.

DEFINITION V.3. A *forbidding grammar* G (abbreviated *N grammar*) is a quadruple $G = (\Sigma, P, S, \Delta)$, where Σ is the *total alphabet of* G, $\Delta \subsetneq \Sigma$ and $\Sigma \setminus \Delta$ are respectively the *terminal* and *non-terminal alphabets of* G, $S \in \Sigma \setminus \Delta$ is the *axiom of* G and P is a finite set of *productions* of the form $A \to w \mid V$, where $A \in \Sigma \setminus \Delta$, $w \in \Sigma^*$ and $V \subseteq \Sigma$.

A word $u \in \Sigma^+$ directly derives a word v according to G, denoted $u \Rightarrow_G v$ if $u = \alpha A \beta$, $\alpha, \beta \in \Sigma^*$, $A \in \Sigma \setminus \Delta$, $v = \alpha w \beta$ and $A \to w \mid V \in P$ and $V \cap alph\ u = \varnothing$. Let $\Rightarrow_G^*$ denote the reflexive and transitive closure of $\Rightarrow_G$. (If no confusion is possible we use $\Rightarrow^*$ and $\Rightarrow$.)

The *language of* G, denoted $L(G)$ is defined by $L(G) = \{v \in \Delta^* : S \Rightarrow^* v\}$. ∎

The class of languages generated by forbidding grammars is denoted $\mathscr{L}(N)$.

LEMMA V.3. *For every N grammar there exists an equivalent symmetric S grammar.*

Proof. Let $G = (\Sigma, P, S, \Delta)$ be a N grammar. We assume that $P = \{P_1, \ldots, P_n\}$ is an ordered set of productions and, if $P_j = a \to \alpha \mid V$ for

$1 \leqslant j \leqslant n$, $a \in \Sigma \setminus \Delta$, $\alpha \in \Sigma^*$ and $V \subseteq \Sigma$ we denote $lh(P_j) = a$, $rh(P_j) = \alpha$ and $N(P_j) = V$.

Let $\Sigma_j = \Sigma \setminus N(P_j)$, for $j = 1,\dots, n$, and let $\dot{\Sigma} = \{a_j : a = lh(P_j),\ 1 \leqslant j \leqslant n\}$ with $\Sigma \cap \dot{\Sigma} = \varnothing$.

Let $\Gamma = \Sigma \cup \dot{\Sigma}$. The finite substitution h, from Γ^* into Γ^*, is defined as follows. If $a = lh(P_j)$ for some $j \in \{1,\dots, n\}$, then $a_j \in h(a)$ and $h(a_j) = rh(P_j)$. If there exists no $j \in \{1,\dots, n\}$, such that $a = lh(P_j)$, then $h(a) = \{a\}$. Let $K = \Sigma^* \bar{\Sigma} \Sigma^* \cup \bigcup_{j=1}^{n} \Sigma_j^* \bar{a}_j \Sigma_j^*$.

Clearly $H = (\Gamma, h, S, \Delta, K)$ is a symmetric S grammar generating $L(G)$ and hence the lemma holds. ∎

LEMMA V.4. *For every symmetric S grammar there exists an equivalent N grammar.*

Proof. Let $H = (\Sigma, h, S, \Delta, K)$ be a symmetric S grammar; $K = \bigcup_{i=1}^{n} X_i^* \bar{Y}_i X_i^*$. We construct an equivalent N grammar, $G = (\Gamma, P, Z, \Delta)$ as follows.

> $\Gamma = \Sigma \cup \dot{\Sigma} \cup \Delta'$, where $\dot{\Sigma} = \{\dot{a} : a \in \Sigma\}$, $\Delta' = \{a' : a \in \Delta\}$ and Σ, $\dot{\Sigma}$ and Δ' are pairwise disjoint.

The set of productions P is defined in the following way.

$P = \{\varphi(a) \to \dot{a} \mid \Delta \cup \dot{\Sigma} \cup \varphi(\Sigma \setminus (X_i \setminus \{a\}))) : a \in Y_i,\ 1 \leqslant i \leqslant n\} \cup$

$\{\dot{a} \to \varphi(\alpha) \mid (\Delta \cup (\dot{\Sigma} \setminus \{\dot{a}\}) \cup \varphi(\Sigma \setminus X_i)) : a \in Y_i,\ 1 \leqslant i \leqslant n \text{ and } \alpha \in h(a)\} \cup$

$\{a' \to a \mid ((\Sigma \setminus \Delta) \cup \dot{\Sigma}) : a \in \Delta\}$, where the homomorphism φ, from Σ^* into Γ^* is defined by $\varphi(a) = a$ if $a \in \Sigma \setminus \Delta$ and $\varphi(a) = a'$ if $a \in \Delta$.

$Z = \varphi(S)$.

The proof of the equivalence of G and H is left to the reader. ∎

THEOREM V.4. $\mathscr{L}(\text{sym S}) = \mathscr{L}(\text{N})$.

THEOREM V.5. $\mathscr{L}(\text{sym S}) \supsetneq \mathscr{L}(ET0L)$.

Proof. This is an immediate consequence of Theorem V.4 and Theorems 1 and 2 of Penttonen (1975). ∎

As for $1C$ grammars, we do not have a characterization of the role of the symmetric restriction imposed on them. It seems however that symmetric $1C$ grammars possess quite strong language generating power as is illustrated by the following.

THEOREM V.6. (i) $\mathscr{L}(\text{sym } 1C) \setminus \mathscr{L}(ET0L) \neq \varnothing$.

(ii) $\mathscr{L}(\text{sym } 1C) \supsetneq \mathscr{L}(E0L)$.

Proof. (i) Let $H=(\Sigma, h, S, \Delta, K)$ be defined as follows. $\Sigma=\Sigma_1 \cup \Sigma_2$, where $\Sigma_1=\{a, b, c, d, F\}$, $\Sigma_2=\{S, A, B, C, D, Q\}$, $\Delta=\{a, b, c, d, Q\}$. $K=\Sigma_1^* \bar{\Sigma}_2^+ \Sigma_1^*$ and h is defined in the following way. $h(S)=AQ$, $h(A)=\{ABC, \; aBC\}$, $h(B)=\{BC, bC\}$, $h(C)=\{C, cD\}$, $h(Q)=\{Q^2\}$, $h(D)=\{d^2\}$, and for all $x \in \Sigma_1$, $h(x)=\{F\}$. It is easy to see that $L(H)=L(G_0)$, where $L(G_0)$ is the language, defined in Ehrenfeucht *et al.* (1980), for which $L(G_0) \in \mathscr{L}(1C) \setminus \mathscr{L}(ET0L)$ holds.

(ii) This follows immediately from (i) and Theorem III.1.(i). ∎

The symmetric restriction imposed on C grammars does not change the class of generated languages.

LEMMA V.5. For every C grammar, there exists an equivalent symmetric C grammar.

Proof. Let $L \subseteq \Delta^*$ be generated by a C grammar. Without loss of generality we assume that L is generated by a right-continuous grammar (Theorem 1 of Ehrenfeucht *et al.* (1980)). From the proofs of Lemmas V.1 and V.2 it follows that there exists a balanced C grammar, generating L, which is constructed from the original right-continuous grammar.

Hence, we assume that L is generated by $H=(\Sigma, h, S, \Delta, K)$ with

$$K=\bigcup_{i=1}^{n} (X_i^* \bar{Y}_i^+ \cup \bar{Y}_i^+ X_i^*).$$

Let $H'=(\Gamma, g, S, \Delta, K')$ be a C grammar, which is defined as follows.

$$\Gamma=\bigcup_{i=1}^{n} (X_i \cup \hat{X}_i \cup Y_i \cup \dot{Y}_i) \cup \Delta, \quad \text{where} \quad \hat{X}_i=\{\hat{a} : a \in X_i\}$$

and

$$\dot{Y}_i=\{\dot{a} : a \in Y_i\}, \quad \text{for} \quad i=1, \ldots, n.$$

The alphabets $\bigcup_{i=1}^n (X_i \cup Y_i)$, $\bigcup_{i=1}^n \hat{X}_i$ and $\bigcup_{i=1}^n \dot{Y}_i$ are pairwise disjoint. The finite substitution g, from Γ^* into Γ^*, is defined in the following way.

For all $a \in \bigcup_{i=1}^{n} X_i$, $\hat{a} \in g(a)$ and $a \in g(\hat{a})$,

for all $a \in \bigcup_{i=1}^{n} Y_i$, $\dot{a} \in g(a)$ and $g(\dot{a})=h(a)$,

for all $a \in \Delta \setminus \bigcup_{i=1}^{n} (X_i \cup Y_i)$, $g(a)=\{a\}$.

$$K' = \bigcup_{i=1}^{n} (X_i^* \bar{Y}_i^+ X_i^* \cup \dot{Y}_i^* \bar{X}_i^+ \dot{Y}_i^* \cup \hat{X}_i^* \bar{\dot{Y}}_i^+ \hat{X}_i^*$$
$$\cup \left(\Gamma \Big\backslash \bigcup_{i=1}^{n} (\hat{X}_i \cup \dot{Y}_i)\right)^* \bar{\hat{X}}_i^+ \left(\Gamma \Big\backslash \bigcup_{i=1}^{n} (\hat{X}_i \cup \dot{Y}_i)\right)^* \Bigg).$$

From the definition of K' it follows immediately that H' is a symmetric C grammar.

The equivalence of H' and H is shown as follows.

Let $w \in (\Gamma \setminus \bigcup_{i=1}^{n} (\hat{X}_i \cup \dot{Y}_i))^*$ and let $w \Rightarrow v$. This implies that (1) $w \in X_j^* Y_j^+ X_j^*$ or (2) $w \in X_j^+$, for some $j \in \{1,..., n\}$, and that $v \in X_j^* \{\hat{a}, \dot{a}: a \in Y_j\}^+ X_j^*$ or $v \in \{\hat{a}, \dot{a}: a \in X_j\}^+$, respectively. If $v \Rightarrow u$, this implies in case (1) that $v \in \dot{Y}_j^* X_j^+ \dot{Y}_j^+$ or $v \in \dot{Y}_j^+$ or $v \in \{\{\hat{a}: a \in Y_j\} \cap \hat{X}_l\}^+$ for some $l \in \{1,..., n\}$. Hence (1.1) $w \in X_j^+ Y_j^+$ and $v \in X_j^+ \dot{Y}_j^+$ or (1.1′) $w \in Y_j^+ X_j^+$ and $v \in \dot{Y}_j^+ X_j^+$ or (1.2) $w \in Y_j^+$, $v \in \dot{Y}_j^+$ and $u \in h(Y_j^+)$, or (1.3) $u = w$, respectively. If $v \Rightarrow u$, this implies in case (2) either (2.1) $v \in \dot{Y}_l^+$ for some $l \in \{1,..., n\}$ and hence $w \in Y_l^+$ and $u \in h(Y_l^+)$ or (2.2) $v \in \hat{X}_j^+$ and hence $w = u$. If we have the case (1.1) that $w \in X_j^+ Y_j^+$, $v \in X_j^+ \dot{Y}_j^+$ and $u \Rightarrow t$, then $u \in \hat{X}_j^+ \dot{Y}_j^+$ and $t \in \hat{X}_j^+ h(Y_j^+)$. Clearly t can only be rewritten as $s = x\gamma$, if $t = \hat{x}\gamma$, where $x \in X_j^+$, $\hat{x} \in \hat{X}_j^+$, $w = xy$, $y \in Y_j^+$ and $\gamma \in h(y)$. Analogously in case (1.1′) it is shown that $w = yx$, $y \in Y_j^+$ and $x \in X_j^+$, and $s = \gamma x$, with $\gamma \in h(y)$.

Hence if for a word $w \in (\Gamma \setminus \bigcup_{i=1}^{n}(\hat{X}_i \cup \dot{Y}_i))^*$ there exists a derivation $w \Rightarrow_{H'}^{+} s$ and $s \in (\Gamma \setminus \bigcup_{i=1}^{n} (\hat{X}_i \cup \dot{Y}_1))^*$, then either $w \Rightarrow_H s$ or $w = s$. From this it follows that $L(H') \subseteq L(H)$.

The proof of the converse inclusion can easily be derived from the above and is left to the reader.

Hence H' and H are equivalent and our statement follows. ∎

THEOREM V.7. $\mathscr{L}(\text{sym } C) = \mathscr{L}(C)$.

Proof. $\mathscr{L}(\text{sym } C) \subseteq \mathscr{L}(C)$ follows from Definition V.1. The converse inclusion was proved in Lemma V.5. ∎

VI. DISCUSSION

The aim of the research presented in this paper was to investigate the intuitive notions of sequential and parallel rewriting in the general framework of selective substitution grammars. The formalization of these notions led to S grammars and L grammars, respectively, which correspond to the "extreme" forms of sequential and parallel rewriting. In order to see the difference between sequential and parallel rewriting in a better perspective we also investigated C grammars which formalize a way of rewriting which is "in between" sequential and parallel. Each of these three classes of grammars was separately investigated and a comparative study of the classes of languages they generate was done.

A number of technical problems should be settled in order to give a more complete picture of the situation. Among these are the following ones.

(1) $\mathscr{L}(1S) = \mathscr{L}(CF)$?

(2) $\mathscr{L}(2S) = \mathscr{L}(S)$?

(3) $\mathscr{L}(2S) \subseteq \mathscr{L}(1C)$?

(4) $\mathscr{L}(ET0L) \setminus \mathscr{L}(1C) = \varnothing$?

(5) Is $\mathscr{L}(C) = \mathscr{L}(S)$ equal to the class of context-sensitive languages?

In addition to solving concrete technical problems like the above one should pursue the investigation of several problem areas within the framework adopted in this paper. For example:

(i) Investigate the role of erasing productions in various classes of grammars, considered in our paper. An example of a concrete question in this direction is: $\mathscr{L}(s - S) = \mathscr{L}(CF)$?

(ii) Investigate the role of the deterministic restriction; a grammar satisfying this restriction has the property that for each letter it has precisely one production to rewrite it.

(iii) Investigate the combinatorial structure of languages considered in this paper.

Acknowledgment

The second author gratefully acknowledges the support of NSF Grant MCS 79-03838.

RECEIVED: October 1, 1980

References

BERSTEL, J. (1979), "Transductions and Context-free Languages," Teubner-Verlag, Stuttgart.

EHRENFEUCHT, A., MAURER, H., AND ROZENBERG, G. (1980), Continuous grammars, *Inform. Contr.* **46**, 71–91.

KLEIJN, H. C. M. AND ROZENBERG, G. (1980), Context-free like restrictions on selective rewriting, *Theoret. Comput. Sci.*, in press.

LOMKOVSKAJA, M. V. (1972), O nekotoryh svojstvah k-uslovnyh grammatik, *Naučno- Tehn. Inform. Ser. 2*, No. 1, 16–21.

PENTTONEN, M. (1975), *ET0L*-grammars and *N*-grammars, *Inform. Process. Lett.* **4**, 11–13.

ROZENBERG, G. (1977), Selective substitution grammars, (towards a framework for rewriting systems). Part I: definitions and examples, *Elektron. Informationsverarb. Kybernet.* **13**, 455–463.

ROZENBERG, G. AND SALOMAA, A. (1980), "The Mathematical Theory of L Systems," Academic Press, New York.

ROZENBERG, G. AND WOOD, D. (1980), Context-free grammars with selective rewriting, *Acta Inform.* **13**, 257–268.

SALOMAA, A. (1973), "Formal Languages," Academic Press, New York.